ACCELEROMETRY MEASUREMENT OF PHYSICAL ACTIVITY AND SEDENTARY BEHAVIOUR IN PRE-SCHOOL CHILDREN

JANE HISLOP

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

QUEEN MARGARET UNIVERSITY

2013

ABSTRACT

This thesis is based on six studies which address questions around the use of accelerometers to measure physical activity and sedentary behaviour of pre-school children: are shorter epochs more accurate? Which epochs are most accurate? Are there advantages to using triaxial accelerometers? Which cut-points are most accurate? Are different generations of Actigraph accelerometers comparable? What is the recommended wear time to provide a reliable estimate of habitual physical activity and sedentary behaviour?

Analysis of 7-10 day accelerometry data, collected from 31 pre-school children (mean (SD) age 5.9 (0.7) y), suggests that shorter epochs (15 s) result in significantly greater estimates of time spent in moderate-to-vigorous physical activity (MVPA) in comparison to 60-s epochs (p <0.05). When compared against a direct observation method, Children's Activity Rating Scale (CARS), with 32 pre-school children (4.4 (0.8) y) during 1 hour of free-play, 15-s epochs were more accurate than 60-s epochs. Comparison of the triaxial RT3 against a uniaxial accelerometer, suggests no advantage of the RT3 accelerometer. The Puyau et al. (2002) cut-points had the 'best' agreement with estimates of sedentary behaviour, light intensity and MVPA against the CARS. Different generations of accelerometers were not comparable, however, application of a correction factor to the GT1M data (7164 = GT1M/0.91) may improve comparability of total physical activity. Finally, analysis of 7 day accelerometry data from 112 pre-school children (3.7 (0.7) y) suggests that 3 days of 7 hours provides a reliable estimate of habitual physical activity and that inclusion of weekend days is not necessary.

This thesis highlights the implications that methodological decisions can have over apparent estimates of physical activity and sedentary behaviour and has made recommendations for accelerometry use. Ideally, there needs to be a move towards consensus, as, only by adopting standardised approaches to accelerometry use, will comparison between study outcomes become meaningful.

Key words: accelerometer, physical activity, pre-school, Actigraph, measurement

ACKNOWLEDGEMENTS

I would like to thank my supervisors Dr Cathy Bulley and Professor Tom Mercer, from the School of Health Sciences, and Professor John Reilly, from Strathclyde University, for their guidance, encouragement and feedback throughout the years. Thank you to Robert Rush from the School of Health Sciences for his statistical advice.

I would like to thank Vicky Penpraze for guidance with the SPARKS study (Chapter 1) and Andrew Grainger for the final study (Chapter 8). I'd also like to thank Andrew and Richard Wilson for their technical support.

I would like to thank the Physiotherapy Research Foundation for funding equipment for this study and time to undertake the data collection.

Thank you to my friends for your support, particularly Marti Balaam and family.

I would like to thank my family in particular, David Gardner for his patient proof reading, Robert for his encouragement to keep going and Cathryn and Alexander who have been the incentive to get this finished.

Finally, I would like to thank the nurseries for their involvement and to the parents and children for their time and input.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

APPENDICES

PUBLICATIONS ARISING FROM THIS THESIS

Journal publications:

Hislop, J.F., Bulley, C., Mercer, T.H. and Reilly, J.J. 2012a. Comparison of accelerometry cut points for physical activity and sedentary behavior in preschool children: a validation study. *Pediatric Exercise Science*, 24 (4) pp.563-576.

Hislop, J.F., Bulley, C., Mercer, T.H. and Reilly, J.J. 2012b. Comparison of epoch and uniaxial versus triaxial accelerometers in the measurement of physical activity in preschool children: a validation study. *Pediatric Exercise Science,* 24 (3) pp.450-460.

Reilly, J.J., Penpraze, V., Hislop, J., Davies, G., Grant, S. and Paton, J.Y. 2008. Objective measurement of physical activity and sedentary behaviour: review with new data. *Archives of Disease in Childhood*, 93 (7) pp.614-619.

Poster presentations:

Hislop, J.F. Bulley, C. Mercer, T. Reilly, J.J. (2013) How many hours and days of data provide reliable estimates of habitual physical activity in preschool children? ICAMPAM conference, Amherst, USA

Hislop, J.F. Bulley, C. Mercer, T. Reilly, J.J. (2011) Concurrent validity of the GT1M with MTI Accelerometer in Preschool children during free-play. World Physiotherapy Congress, Amsterdam, Netherlands.

Hislop, J.F. Bulley, C. Mercer, T. Reilly, J.J. (2011) Reliability and concurrent validity of GT1M and MTI Actigraph Accelerometers using a mechanical setup. ICAMPAM conference, Glasgow, Scotland.

Hislop, J.F, Bulley, C. Mercer, T. Reilly, J.J.(2009) ‘Physical Activity of Pre-school children in Lothian. Phase 1: Study of Accelerometer Methods’. Poster presentation at staff Conference, Queen Margaret University, Edinburgh.

ABBREVIATIONS

AEE	Activity energy expenditure
AP	Anteroposterior
CARS	Children's Activity Rating Scale
cpm	Counts per minute
dm	Mean difference
EE	Energy expenditure
BMI	Body Mass Index
BMR	Basal metabolic rate
DIT	Diet induced thermogenesis
HR	Heart rate
LOA	Limits of agreement
LPA	Light physical activity
M	Mean
Mdn	Median
MET	Metabolic equivalent units
ML	Mediolateral
MPA	Moderate physical activity
MVPA	Moderate-to-vigorous intensity physical activity
PE	Physical education
r_p	Pearson's correlation
r_s	Spearman's correlation
ROC	Receiver Operating Characteristic Curve
RMR	Resting metabolic rate
SD	Standard deviation
Sed	Sedentary behaviour
TPA	Total physical activity
TEE	Total energy expenditure

VPA	Vigorous physical activity
VO_2	Oxygen uptake

CHAPTER 1 : GENERAL INTRODUCTION AND LITERATURE REVIEW

1.1 INTRODUCTION

The focus of this thesis is to address methodological questions relating to the accurate use of accelerometers in objectively measuring free-play in pre-school children. The purpose of this chapter is:

- To introduce the key terms used within the thesis;
- To outline the background for the thesis;
- To critically review the literature relating to measurement of physical activity;
- To provide a rationale for the thesis by identifying the key methodological questions relating to accelerometry use in pre-school children on which the thesis is based;
- To outline the structure of the thesis;
- To state the aims of the thesis.

1.2 SEARCH STRATEGIES

To identify relevant literature, a search was carried out using MEDLINE, EMBASE, CINHAL, and SPORTDiscuss databases, from 1985 to 2012, using the key words: 'young child' or 'pre*school*' or 'nursery' or 'kindergar*' or 'early childhood' in combination with 'physical activity' or 'movement' or 'sedentary behav* or 'play'. For the second section on the background to the thesis the search was limited to studies between 1992 to 2012 and used the above search terms in combination with 'measure*' or 'direct observation' or 'doubly labeled water' or 'DLW' or 'calirometry' or 'heart rate' or 'pedomet*' or 'acceleromet*'. In the final section on measurement of physical activity the previous search terms were used in combination with the terms 'validity' or 'reliability' or 'calibration' to identify appropriate studies.

1.3 DEFINITION OF KEY TERMS

1.3.1 Definition of physical activity, sedentary behaviour, exercise and fitness in young children

The terms ‘physical activity’, ‘exercise’ and ‘fitness’ are often used interchangeably, however, these are distinct constructs. One of the earliest and most frequently cited definitions of physical activity is by Caspersen et al. (1985):

> Physical activity is defined as any bodily movement produced by skeletal muscles which results in energy expenditure (Caspersen et al. 1985, p. 126).

Physical activity is a complex multi-dimensional behaviour, which is often categorised and measured using the variables frequency, intensity, duration and type (Valanou et al. 2006). The type of physical activity relates to the activity behaviour, such as walking, sitting, standing or lying. The duration and frequency of physical activity can be measured by time and number of bouts of physical activity over a period of time respectively, while intensity can be measured by level of effort or rate of energy expenditure (EE) (Department of Health, Physical Activity, Health Improvement and Protection 2011).

When intensity of physical activity is expressed in terms of rates of energy expenditure (EE) it is in the form of metabolic equivalent units (METs). The MET values are based on multiples of resting metabolic rate (RMR) which can be estimated or measured directly (Ridley and Olds 2008). One MET is the energy cost measured by baseline oxygen consumption (VO_2), which for adults is taken to be 3.5 $ml \cdot kg^{-1} \cdot min^{-1}$ (Ridley and Olds 2008). For children it is more complex as their RMR is substantially higher than adults and varies according to the child’s age, sex, body mass and pubertal status (Harrell et al. 2005; Ridley and Olds 2008). It is estimated that at 5 years of age the RMR for children is as high as approximately 6 $ml \cdot kg^{-1} \cdot min^{-1}$ and declines to adult levels by 18 years of age (Schofield 1985). As a consequence, adult baseline RMR values are not suitable for estimating MET values for children (Ridley and Olds 2008).

Conventionally, the intensity of physical activity is categorised into light, moderate and vigorous. Moderate and vigorous intensity physical activities are often grouped together into one category of moderate-to-vigorous intensity physical activity (MVPA). The American College of Sports Medicine for Physical Activity have defined adult MET thresholds which relate to different intensities of physical activity. These are: 1.5 to < 3 METs for light intensity physical activity (LPA); 3 to < 6 METs for MVPA; 6 to 9 METs for hard intensity physical activity and > 9 METs for very hard intensity physical activity (Freedson et al. 1998; Hendelman et al. 2000; Swartz et al. 2000).

Light intensity physical activity usually refers to activities such as slow walking, while moderate intensity physical activity will raise the heart rate and leaves a person feeling out of breath, and includes activities such as brisk walking (Cavill et al. 2006). Vigorous intensity physical activity includes activities such as running (Freedson et al. 1998; Kozey et al. 2010a). However, estimating MET values for children is not straightforward as their energy cost for undertaking activities is higher than adults, possibly due to developmental reasons, for example having less efficient gait patterns or faster respiration rates (Rowland and Green 1988). Ridley et al. (2008) has synthesised the evidence relating to energy costs for a range of activities for school-aged children and adolescents (6 to 18 y). However, there is limited evidence of the appropriate MET values for use with young children.

It should be noted that while energy expenditure (EE) results from bodily movement, physical activity is only one factor which contributes to EE (Oliver et al. 2007b). An individual's total energy expenditure (TEE) has three components: resting energy expenditure (REE), diet-induced thermogenesis (DIT) and activity energy expenditure (AEE) (Muller and Bosy-Westphal 2003). AEE is the most variable component of TEE (Westerterp and Kester 2003). REE is influenced by an individual's RMR which is the minimal energy expenditure required to fuel basic physiological functions (e.g. heartbeat, respiration, muscle contractions) (Goran and Treuth 2001). RMR is variable between individuals with 80-90% of the variation accounted for by differences in an individual's organ and muscle mass, known as fat free mass (FFM), their fat mass, age, gender, ethnicity and level of physical activity (Goran and Treuth 2001; Kozey-Keadle et al. 2010). In summary, EE is a complex variable determined by a variety of factors such that two individuals may expend different amounts of energy for the same physical activity depending on their individual physical and physiological characteristics.

Sedentary behaviour is argued to be a separate behavioural construct from physical activity with different determinants and the two constructs should not be regarded as opposite sides of the same coin (Pate et al. 2011). Sedentary behaviour is not simply the absence of physical activity but predominantly involves sitting and lying down activities (Pate et al. 2008).

The term sedentary behaviour has been defined as a behaviour that produces no movement and which results in minimal (or no) physiological gain (Pettee Gabriel et al. 2012). It has also been used to refer to situations where there is no trunk translocation or energy expenditure is less than 1.5 METs (Martin et al. 2011). The most recent definition states that sedentary behaviour is any waking activity which involves energy expenditure of less than or equal to 1.5 METs and which is undertaken in a seated or reclined position (Sedentary Behaviour Network, 2012).

Exercise is described as a sub-category of physical activity and is defined as:

> …a subset of physical activity that is planned, structured and repetitive and has a final or an intermediate objective the improvement or maintenance of physical fitness. (Caspersen et al. 1985, p. 126)

As a consequence physical activity is defined as being 'exercise' when the intention is to maintain or promote fitness (Trost 2001).

Physical fitness is defined as 'a set of attributes that are either health- or skill-related' (Caspersen et al. 1985, p. 126). While physical activity and exercise relate to the movements undertaken by the body, fitness is a state characterised by certain traits and associated with factors such as low risk of premature development of chronic disease (Baranowski et al. 1992). Physical activity and exercise are the processes that may assist in attaining an outcome of physical fitness (Rice and Howell 2000).

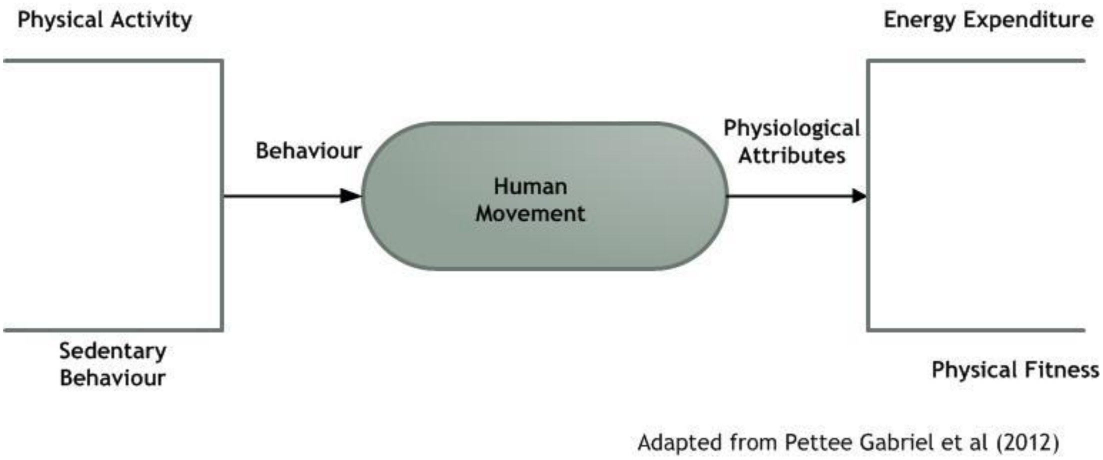
Bar-Or and Malina (1995) argue that in young children the focus should be on developing motor skills rather than 'fitness' and only once children have achieved motor proficiency (around the age of 10 years) should fitness begin to be emphasised.

Recently there have been calls to establish more precise terminology around the constructs of physical activity as this is fundamental for consistency of interpretation when measuring physical activity (Pettee Gabriel et al. 2012). Pettee Gabriel et al. (2012) revisited the earlier Caspersen et al. (1985) definitions to propose a conceptual framework for physical activity. The framework (Figure 1.1) takes into account sedentary behaviour and differentiates between the behavioural and physiological elements of human movement. Pettee Gabriel (2012) defined physical activity as:

> behavior that involves human movement, resulting in physiological attributes including increased energy expenditure and improved physical fitness. (Pettee Gabriel et al. 2012, p. S15)

Figure 1.1 presents the conceptual framework for physical activity proposed by Pettee Gabriel et al. (2012)

Figure 1.1: Conceptual framework for physical activity.



This thesis will focus on the use of accelerometers to accurately quantify the behavioural elements of physical activity i.e. the intensity, duration and frequency of physical activity and the frequency and duration of sedentary behaviour.

1.3.2 Definition of free-play

Unstructured active play is argued as being the principle mode of physical activity in young children (Dwyer et al. 2009) and much of the research into younger children's physical activity make use of the term 'play'. Play can be defined as:

> …spontaneous activity in which children engage to amuse and occupy themselves (Burdette and Whitaker 2005, p. 46).

Burdette and Whitaker (2005) reason that 'play' is a more appropriate term to use in the promotion of movement in young children whose physical activity patterns are different from older children. It has been argued that while 'play' can appear to be 'purposeless' it has an important role in functionality not only for physical development but for the development of cognitive and social skills (Burdette and Whitaker 2005; Pellegrini and

Smith 1998). This is in contrast to ‘exercise’, which is structured activity that is repetitive and purposeful (Caspersen et al. 1985) and tends to be adopted in older children and adults.

Pellegrini and Smith (1998) use the term ‘physical activity play’ which they argue has the distinguishing features of physical activity which is undertaken in a playful context at an intensity where metabolic activity is well above resting levels, i.e. at moderate-to-vigorous levels of intensity. In contrast, Timmons et al. (2007) argue that ‘physical activity play’ can be undertaken at various levels of intensity. Findings of observational research of young children supports the idea that young children’s physical activity behaviour is characterised by varying intensity, where short bursts of high intensity activity are interspersed with longer periods of low intensity activity and rest (Bailey et al. 1995; Oliver et al. 2009).

In young children, physical activity is achieved predominantly through active play and this becomes less so in school-aged children (Dwyer et al. 2009). According to Dwyer et al. (2009) if physical activity is to be understood, promoted and measured in young children then this needs to be within the context of active play.

In the USA, one of the first nationally published guidelines for young children’s physical activity, the Active Start Guidelines from the National Association for Sports and Physical Education (NASPE), state that:

> Preschoolers should accumulate at least 60 minutes of structured and at least 60 minutes and up to several hours of unstructured physical activity per day. (National Association for Sport and Physical Education 2002, p. 9)

In these guidelines ‘structured’ activity refers to organised and planned movement experiences to help children develop fundamental movement skills (FMS). Unstructured activity refers to the intermittent movement experiences that may occur through ‘free-play’. FMS represent the basic performance competencies that underlie and are required for carrying out complex motor skills involved in different types of physical activity (Payne and Isaacs 1995). There are two categories of FMS: locomotor skills (e.g. run, hop and jump) and object-control skills (e.g. catch, throw and kick) (Cliff et al. 2009a).

In the current thesis the term ‘free-play’ is used to refer to unstructured physical activity play, which is child-led and includes all intensities of physical activity. As the current thesis focuses on the accuracy of accelerometers to measure free-play, this refers to physical activity which involves exercise play and rough and tumble play, as opposed to ‘games’, such as football or tennis, which more applies to the type of structured play undertaken by older children (Pellegrini 2010).

1.3.3 Definition of pre-school children

The terms ‘young children’, ‘pre-school’, ‘early childhood’ and ‘early years’ are used interchangeably in the literature to refer to children less than 5 years of age (Oliver et al. 2007b). The ‘early years’ are defined as being from birth to 5 years of age and include three developmental stages, each with distinctive physical activity patterns (Cliff et al. 2009a). ‘Infant stage’ is described as from birth to 12 months during which the child develops basic motor skills. ‘Toddler stage’ is from 1 to 3 years, where the child develops locomotor skills, such as jumping, hopping, skipping and running as well as object control through kicking and catching a ball. Finally, pre-school age, from 3 to 5 years of age, is characterised by further progression of the locomotor and object-control skills developed during the toddler stage (Cliff et al. 2009a). This thesis is concerned with children in the pre-school period between the ages of 3 to 5 years, who will also be referred to as young children.

1.3.4 Summary

The key points from this section are:

- Physical activity and sedentary behaviour are different behavioural constructs;
- Physical activity can be considered as a behaviour which has dimensions of type, intensity, frequency and duration;
- Energy expenditure, which is a physiological response to physical activity, is also a different construct from physical activity;
- Energy expenditure is a complex variable and use of MET values to estimate intensity of physical activity in young children is not straightforward;
- In this thesis the term pre-school will refer to children aged 3 to 5 years of age and free-play will refer to the unstructured, intermittent free-living activity believed to be typical of young children.

1.4 BACKGROUND TO CURRENT THESIS

1.4.1 Physical activity and health

Studies to date have reported low and declining levels of physical activity (Basterfield et al. 2011) and high levels of sedentary behaviour in young children (Hinkley et al. 2012c; Kelly et al. 2007). These reports are of concern as low levels of physical activity and increased time spent in sedentary behaviour in children have been linked to chronic adult disease risk factors including cardiovascular risk factors (Andersen et al. 2006; Janssen and LeBlanc 2010; Strong et al. 2005; Tremblay et al. 2011), increased adiposity in childhood (Steele et al. 2009) as well as being independent risk factors for metabolic syndrome (MeS) (Brage et al. 2004a; Ekelund et al. 2006; Ekelund et al. 2009). MeS refers to the clustering of risk factors for coronary heart disease, stroke and type II diabetes (Lambourne and Donnelly 2011). Furthermore it is argued that the risk factors for chronic disease and MeS as a consequence of sedentary lifestyles may be present in young children (Saakslahti et al. 2004; Sirard and Pate 2001). Declining levels of physical activity have also been attributed to the dramatic increase in childhood obesity in the UK since the 1980s (Parsons et al. 1999) and while levels of obesity in child and adult populations have plateaued in recent years, it is maintained that resolving the childhood obesity ‘epidemic’ should remain a public health priority (Rokholm et al. 2010).

It is argued that as physical activity is one of the modifiable risk factors for MeS, that prevention should not only target childhood obesity but also aim to promote physical activity (Ekelund et al. 2009). Although not strong, there is some evidence to suggest that increasing physical activity levels in children could reduce the risk of chronic disease by improving a child’s cardio-metabolic profile (Janssen and LeBlanc 2010). In addition increasing physical activity may contribute to the prevention of obesity in young children (Guinhouya et al. 2011) as well as enhancing bone health for young children from engagement in high intensity physical activity (Janz et al. 2010).

The pre-school period (3 to 5 years of age) is argued to be one of the critical periods of childhood during which the long term regulation of energy balance may be programmed (Dietz 1997). This, together with the fact that lifestyle behaviours are thought to track from pre-school to childhood, and subsequently into adulthood (Biddle et al. 2010; Malina 1996) means that the early years may be a critical time for promoting physical activity and preventing sedentary habits developing (Goldfield et al. 2012).

However, the relationship between physical activity, sedentary behaviour and health in the early years is not fully understood and there is a call for more robust studies to explore these relationships and to determine whether strategies for increasing physical activity could lead to improved health indicators (Guinhouya et al. 2011). In addition, there is limited evidence for the frequency, duration, intensity and type of physical activity which is associated with health indicators, or which is necessary for health gain in young children, and further research is needed (Timmons et al. 2012).

Critical to the research which seeks to gain further insight into the relationship between physical activity, sedentary behaviour and health in the early years, is accurate methods of measuring the constructs of physical activity. In addition, agreement on the approaches to measuring physical activity will allow better understanding of these relationships (Timmons et al. 2012).

There are a few studies to date which have gathered population-based objective data describing levels and patterns of physical activity and sedentary behaviour in young children (Riddoch et al. 2007). The studies which have been conducted suggest that pre-school children exhibit high levels of sedentary behaviour (Hinkley et al. 2012c; Jackson et al. 2003) with low levels of MVPA and with only a small number of studies suggesting pre-school children are reaching recommended levels of physical activity per day (Cardon and Bourdeaudhuij 2008). However, the accuracy of population-based studies has been called into question; either for their use of methods or their interpretation of findings in such a way which may not accurately capture the physical activity and sedentary behaviour of young children (Rowlands and Eston 2007).

There has been much debate in the literature into whether young children are sufficiently active for health (Tucker 2008) and conflicting conclusions have been reached. The following section will outline physical activity recommendations and those for sedentary behaviour and will review the evidence as to whether young children are meeting recommendations for health.

1.4.2 Recommendations and guidelines

In this section the terms ‘guidelines’ and ‘recommendations’ will be used interchangeably to refer to the advice on physical activity necessary for health gain. It is only in relatively recent years that physical activity guidelines have been developed for children, as it was previously assumed that children were sufficiently active and therefore guidelines were not required (Katzmarzyk and Ardern 2004). Prior to 2009, the USA was the only country to have developed guidelines specific to younger children (3 to 5 years) (National Association for Sport and Physical Education 2002). Until 2011, the Scottish guidelines for children's physical activity were applied to all ages groups of children (3 to 18 years) (Scottish Executive 2003). In recent years the need for age-specific guidelines were felt to be important given the physiological and developmental differences between children of different age groups e.g. between toddlers, pre-schoolers, school-aged children and adolescents (Skouteris et al. 2012).

In 2011, specific guidelines for younger children were jointly published by the health departments for Scotland, England, Wales and Northern Ireland (Department of Health, Physical Activity, Health Improvement and Protection 2011). These recommendations were based on the 'best' evidence to date, although as discussed earlier, the evidence of the exact dose-response relationship between physical activity and health in the pre-school population is limited (Timmons et al. 2012). There are now four countries; United Kingdom, USA, Australia and Canada, that have published specific recommendations for the pre-school age-group. Guidelines for sedentary behaviour have also been published. The recommendations for physical activity are outlined in Table 1.1 and in Table 1.2. These tables summarise physical activity and sedentary behaviour recommendations which were used for young children and published prior to 2009 (Table 1.1) and those published between 2009 and 2012 (Table 1.2).

Table 1.1: Physical activity guidelines/recommendations for children by country prior to 2009.

Country	Organisation	Physical Activity	Sedentary Behaviour	Specific to young children?
UK	Scottish Executive (2003)	Children should accumulate ≥ 60 min/day of MVPA.	NA	No
	UK Health Education Authority (Biddle et al. 1998)	60 min/day of at least moderate intensity activity, or, for those doing little activity currently, at least 30 min of at least moderate intensity activity, plus activities to enhance muscular strength, flexibility and bone health at least 2 times/week.	NA	No
USA	National Association for Sport and Physical Education (2002)	Children (3 - 5 y) should accumulate at least ≥ 60 min/day of structured and at least ≥ 60 min/day and up to several hours per day of unstructured activity (i.e. ≥120 min/day of total physical activity).	Should not be sedentary for more than 60 min at a time, except when sleeping.	Yes
	American Academy of Pediatrics (2006)	Pre-school aged children (4 - 6 y) free-play should be encouraged with emphasis on fun.	Reduce sedentary time spent in transport (car) or stroller, and no more than 2 hour/day screen time.	Yes
Australia	Australian Department of Health and Aging (Department of Health and Ageing 1999; Pate et al. 1999)	30 min/day of moderate activity on most, preferably all days of the week, plus vigorous exercise for 30 min/day 3 - 4 times/week.	NA	No
Canada	Health Canada (2002)	Children and youths should increase the time they currently engage in moderate or vigorous physical activity by at least 30 min/day (in periods of 5-10 min) progressing to ≥ 90 min/day or more of increasing vigorous physical activity.	Decrease time spent in sedentary activity (television, video games, internet) by at least 30 min/day, eventually decreasing by ≥ 90 min/day the amount of time spent daily on these activities.	No

MVPA: moderate-to-vigorous physical activity; NA: not available.

Table 1.2: Physical activity guidelines/ recommendations for young children by country 2009-2012.

Country	Author	Physical Activity	Sedentary Behaviour	Specific to young children?
Australia	Australian Government Department of Health and Aging (2009)	Pre-schoolers (3 - 5 y) should be physically active every day for at least 3 hour/day spread throughout the day. This can include light activity like standing up and moving around and playing as well as more vigorous activity like running and jumping.	Should not be sedentary, restrained, or kept inactive for more than 1 hour at a time, with the exception of sleeping. No more than 1 hour/day screen time.	Yes
Canada	Canadian Society for exercise physiology (2012)	Pre-schoolers (3 - 4 y) should accumulate at least 180 min/day of physical activity at any intensity spread throughout the day. Include a variety of activities in different environments, activities that develop movement skills; progression toward at least 60 min/day of energetic play by 5 years of age.	Minimise time pre-schoolers (3 - 4 y) spend being sedentary during waking hours. This includes prolonged sitting or being restrained (e.g. stroller, high chair) for more than 1 hour. For children 2 - 4 y, screen time should be limited to 1 hour/day; less is better.	Yes
UK	Department of Health, Physical Activity, Health Improvement and Protection (2011)	Children of pre-school age (under 5's who are capable of walking unaided who have not started school) should be physically active daily for at least 180 min/day (3 hour/day), spread throughout the day.	Minimise amount of time being sedentary (being restrained or sitting) for extended periods (except time for sleeping).	Yes
USA	National Association for Sport and Physical Education (2009)	[1]Children (0 - 5 y) should accumulate at least 60 min/day of structured and several hours of unstructured activity per day.	Should not be sedentary for more than 60 min at a time, except when sleeping.	Yes
	Institute of Medicine and National Academies(2011)	Child care providers should provide opportunities for children to be physically active during the day. 15 min/hour of light, moderate and vigorous activity in childcare. Community and built environment should promote PA.	Child care providers should limit time pre-schoolers spent in sitting and standing still to < 30 min at a time.	Yes
Ireland	Department of Health and Children (2009)	Children (2 - 18 y) should accumulate at least 60 min/day of MVPA, including muscle strengthening, flexibility and bone strengthening exercises 3 times a week.	NA	No

MVPA: Moderate-to-vigorous activity; NA: not available.

1.4.3 Levels of physical activity and sedentary behaviour in pre-school children

The recent UK recommendations state that pre-school children should undertake 180 minutes (3 hours) of daily physical activity which should include both light and energetic activities, such as running, swimming and skipping (Department of Health, Physical Activity, Health Improvement and Protection 2011). The focus of these guidelines is for young children to achieve a daily amount of physical activity rather than being concerned with the intensity of that physical activity (Department of Health, Physical Activity, Health Improvement and Protection 2011). Since 2011, the UK guidance has also included recommendations for sedentary behaviour for young children. While the UK recommendations state that sedentary time should be minimised, the Australian, Canadian and USA recommendation set specific time guidance for sedentary time and for TV or screen time (Australian Government, Department of Health and Ageing 2009; Canadian Society of Exercise Physiology 2012; National Association for Sport and Physical Education 2009; National Institute for Health and Clinical Excellence 2009).

Prior to 2011, the physical activity guidelines in Scotland were for children of all ages to engage in at least 60 minutes of moderate-to-vigorous activity (MVPA) per day (Biddle et al. 1998; Scottish Executive 2003). While the recent UK guidelines have shifted the focus away from a concern with intensity of physical activity, MVPA is still an essential behaviour for health in the pre-school years. In particular MVPA is important for young children for the development of cardiorespiratory systems, for bone health (Janz et al. 2010; Saakslahti et al. 2004) and for influencing obesity risk (Lanigan et al. 2010). Recent research of slightly older children (7 to 9 year olds) suggests that MVPA may be the most valuable behavioural target for obesity prevention (Basterfield et al. 2012).

To review whether young children are meeting recommendations for health Appendix I (Appendix Table I.i) presents a summary of empirical studies published between 1992 and 2012 that have measured ‘habitual’ physical activity and sedentary behaviour in pre-school children. The studies reviewed have made use of a range of subjective and objective approaches to measuring physical activity and sedentary behaviour. The merit of these approaches for use with pre-school children will be discussed in detail in section 1.5 of this chapter. To reflect ‘typical’ or ‘habitual’ physical activity of pre-schoolers those studies which only included part of a day e.g. only time spent in or out of nursery/pre-school were excluded.

For each of the studies reported in Appendix I (Appendix Table I.i) it has been noted in the final column whether it is ‘likely’ that the majority of participants of the study were sufficiently active (i.e. meeting the earlier recommendations of ≥ 60 min/day of MVPA or ≥ 180 min/day of total physical activity) or engaging in more than 60 or 120 min/day of screen time, based on the mean values reported. Given that the raw data for subjects were not available in many studies the mean had to be used and therefore it is an approximation of whether targets are being reached (Tucker 2008). Some studies report the mean percentage of time spent in each activity rather than minutes per day. This has been extrapolated using the mean hours per day (if stated) or using the threshold of more than 7.1% of total time in MVPA, which corresponds to more than 60 minutes over a 14-hour day or more than 21.4% of total physical activity in total time, which corresponds to more than 180 minutes over a 14-hour day. These are conservative estimates given that many studies made use of the criteria of 6 hours of activity per day with a mean wear time reported being close to 10 hours.

A total of 72 studies (published between 1992 and 2012) representing 20,942 participants from 17 countries were reviewed. Of these studies the majority were conducted in the USA (n = 27), with the second most frequently being conducted in Scotland (n = 10) and Australia (n = 10). Physical activity was measured by accelerometers in 60 (83%) of studies and of these 46 (77%) made use of the Actigraph (Fort Walton Beach, FL, USA) uniaxial accelerometers. The remaining studies used other models of accelerometer (n = 15), direct observation (n = 4), pedometry (n = 3), questionnaires/proxy report (n = 3) or heart rate monitoring (n = 4) to measure physical activity.

Forty-four (61%) of the studies reviewed provide information on whether pre-school children accumulated 60 minutes of MVPA per day. The finding of just under half of the studies (n = 20, 45%) suggest that pre-school children achieve considerably less than 60 minutes per day (Alhassan et al. 2007; Cardon and Bourdeaudhuij 2008; Cardon and De Bourdeaudhuij 2007; Cliff et al. 2009a; Dolinsky et al. 2011; Fisher et al. 2005a; Fisher et al. 2005b; Hinkley et al. 2012c; Janz et al. 2010; Kelly et al. 2005; Kelly et al. 2006; Kelly et al. 2007; Montgomery et al. 2004; O'Dwyer et al. 2011; Reilly et al. 2004; Reilly et al. 2006b; Spittaels et al. 2012; Taylor et al. 2009; Van Cauwenberghe et al. 2012; Wen et al. 2010). This is supported by the findings of a meta-analysis of pooled data from 14 studies involving 20,871 children (4 - 18 y) undertaken by Ekelund et al. (2012). In their study Ekelund et al. (2012) found that children accumulated a mean (SD) of 30 (21) min/day of MVPA.

However, in contrast 20 studies reviewed reported that pre-school children undertake more than 60 min/day of MVPA (Cliff et al. 2011; Collings et al. 2013; Cox et al. 2012; Duncan et al. 2008; Finn and Ullmann 2004; Gabel et al. 2011; Heelan and Eisenmann 2006; Hume et al. 2012; Jago et al. 2005; Janz et al. 2002; Janz et al. 2004; Martinez-Gomez et al. 2009; Metallinos-Katsaras et al. 2007; Obeid et al. 2011; Pfeiffer et al. 2009; Sarzynski et al. 2010; Sigmund et al. 2009; Tanaka and Tanaka 2009; Telford et al. 2005; Vale et al. 2010; Williams et al. 2008). One study reported that children achieved 60 min/day during the weekdays but not at weekends. (Denham-Deal 2005). Another study reported that children appear to be either meeting or not meeting the recommendations depending on how the data were analysed (Beets et al. 2011a).

Eleven studies have explored whether children engaged in more than the recommended screen time outlined in the guidance for Australia, Canada and USA (American Academy of Pediatrics 2001; Australian Government, Department of Health and Ageing 2009; Canadian Society of Exercise Physiology 2012). Again there were mixed findings with children reported to being exposed to more than the recommended ≥ 2 hours of screen time per day in five studies (Hume et al. 2012) or having less than 2 hours of screen time in five studies (Hinkley et al. 2012c; Taylor et al. 2009). In the further study there were mixed findings on the amount of screen time depending on the nationality of the parents (Bürgi et al. 2010). In these studies the data collected on the child's screen time relied on proxy reporting by parents which, as a method, is limited by subjectivity and the risk of bias as a consequence of parents providing socially desirable responses (Bringolf-Isler et al. 2012). The limitations of proxy reporting will be considered further in section 1.5.2. While some of the guidelines refer to children not being sedentary for more than 1 hour at a time (Australian Government, Department of Health and Ageing 2009; Canadian Society of Exercise Physiology 2012; National Association for Sport and Physical Education 2009) or for 30 minutes while in care (Institute of Medicine of the National Academies 2011) the studies have reported on total sedentary time per day or per hour. It is not possible to determine if the episodes of children's sedentary behaviour in pre-school are less than 1 hour or not.

Given the release in recent years of the recommendations in the UK, Canada and Australia for pre-school children to accumulate at least 180 minutes of total physical activity per day there are only 12 studies which have explored this in the pre-school population. Of these, seven studies report that pre-school children are achieving this recommendation of 180 minutes per day (Heelan and Eisenmann 2006; Janz et al. 2002; Janz et al. 2004; Martinez-

Gomez et al. 2009; Metallinos-Katsaras et al. 2007; Obeid et al. 2011; Telford et al. 2005) and four studies report that children are not meeting this recommendation (Alhassan et al. 2007; Cliff et al. 2009a; Hinkley et al. 2012c; Taylor et al. 2009).

While these conflicting findings are surprising they could in part be explained by the different measurement approaches adopted (e.g. self-report, accelerometry, heart rate etc.) within studies which makes cross-comparison difficult. In addition, even in studies which use the same methods such as accelerometry, different approaches to data collection, processing and analysis have been adopted. These different approaches are a reflection of the lack of consensus in the literature on the optimal approach to use with methods such as accelerometry (Cliff et al. 2009b). The lack of standardisation between the approaches adopted in studies means that conflicting estimates of time spent in physical activity and sedentary behaviour have been reported. As a consequence of this methodological uncertainty there is confusion over whether pre-school children are sufficiently active for health and the true characteristics of physical activity in pre-school children are still not well established (Bornstein et al. 2011).

1.4.4 Summary

Many of the findings of the studies reviewed appear to contradict each other. Whether pre-school children are sufficiently active for health, either achieving 60 min/day of MVPA or 180 min/day of total physical activity is not conclusive. There are also limited number of studies to date which have explored sedentary behaviour and most of these studies rely on subjective methods to measure sedentary behaviour which, as will be discussed in the next section of this review, have limitations.

Several factors could be attributed to the inconsistencies in findings between studies. In particular, these inconsistencies may be an artefact of the different measurement methods adopted and the variety of different methodological decisions used between studies. It is argued that there is a need for consensus on the methodological decisions which influence accurate quantification of physical activity in young children (Cliff et al. 2009b; Ojiambo et al. 2011). The following section will consider in greater depth the approaches to measuring physical activity and sedentary behaviour and their suitability for use with pre-school populations.

1.5 MEASUREMENT OF YOUNG CHILDREN'S PHYSICAL ACTIVITY

1.5.1 Introduction

As discussed, young children's physical activity behaviour is believed to be characterised by intermittent and sporadic bursts of vigorous activity (Bailey et al. 1995; Baquet et al. 2007). As a result, accurate means of quantifying physical activity in this population is particularly challenging, and requires methods that are able to capture the intermittent nature of young children's physical activity (Loprinzi and Cardinal 2011).

Sirard and Pate (2001) describe three categories of methods to measure physical activity: primary, secondary and subjective methods. Primary methods refer to the criterion methods which provide an exact measure of physical activity. However, this is argued as being problematic as currently there is no 'gold standard' method which can accurately measure all aspects of physical activity of pre-school children (Oliver et al. 2007b). There are however criterion methods for measuring energy expenditure, such as doubly labeled water (DLW) and indirect calorimetry and behavioural criterion methods, such as direct observation, against which other methods have been validated. Secondary methods include objective measurement techniques such as heart rate, pedometry and accelerometry, which provide an objective assessment of physical activity (Sirard and Pate 2001). Finally, subjective methods include methods such as self-report or proxy report questionnaires. In the following section the methods applicable to each category will be critically discussed. Consideration will be given to the validity and reliability of the methods as well as to the suitability for the measurement of habitual physical activity behaviour in pre-school children.

The validity of a measurement instrument refers to the extent to which it measures what it set out to measure (Terwee et al. 2010). There are different types of validity such as criterion and concurrent validity (Terwee et al. 2010). Criterion validity refers to the relationship between a measure and the 'gold standard' or criterion measure (Terwee et al. 2010). Concurrent validity examines the agreement between two measures of unknown validity for example between different movement sensors such as accelerometers and pedometers (Lubans et al. 2011).

Reliability refers to the consistency or agreement between multiple measures, either by different observers (or inter-rater reliability) or with the same observer (intra-rater reliability) (Lubans et al. 2011).

To interpret the validity and reliability from studies, de Vries (2009) has provided useful criteria. They rated the criterion validity as 'good' if the correlation coefficient was ≥ 0.75 and the construct validity as 'good' if the correlation coefficient was ≥ 0.60 (de Vries et al. 2006). For reliability intraclass correlation coefficients (ICCs) were argued to be a satisfactory measure of reliability with values of ≥ 0.70 suggesting good reliability.

1.5.2 Subjective methods of measuring physical activity

Subjective methods of measuring physical activity include questionnaires and self-report methods of one's own behaviour or that of others (either parent proxy or teacher reports). Questionnaires have the advantage of being inexpensive and non-invasive and for these reasons they are often used in large-scale population surveys (Oliver et al. 2007b). They can also be used to provide important contextual information on behaviours such as time spent watching television or engaging in electronic media (Dwyer et al. 2011). Unfortunately however, the inability of young children to accurately recall the intensity, duration and frequency of physical activity (Welk et al. 2000b) means that the use of self-report methods are not recommended in children under the age of 10 years (Sallis 1991).

One of the main limitations of proxy questionnaires, where-by parents or carers complete a questionnaire on behalf of the child, is the potential for inaccuracies as a consequence of reporting bias (Noland et al. 1990). This may in part be due to socially desirable responses being given (Bringolf-Isler et al. 2012). Similar to studies of older children, time spent in sedentary activity has been found to be under-reported by parents when compared against objective methods (e.g. accelerometry) and time spent in MVPA tends to be overestimated by parents (Bender et al. 2005). This systematic error has been found to become larger as the magnitude of reported time increases (Anderson et al. 2005; Corder et al. 2009). In addition, Norland et al. (1990) found that parents consistently rated their children's activity levels higher than teachers.

Several proxy questionnaires have been developed and validated with pre-school children. However, many of the validation studies only investigate concurrent validity as they compare questionnaires against secondary measures, such as accelerometers, pedometers and heart rate, which have unknown validity (Lubans et al. 2011). As there are several outstanding methodological questions with secondary measures, such as how accelerometry data should be processed and analysed or the influence of emotion on heart rate, this limits their ability to act as true criterion methods. A summary of validation studies of proxy questionnaires for pre-school children are presented in Table 1.3.

Table 1.3: Summary of validation studies of proxy questionnaires with young children.

Details of questionnaire	Parent proxy (PP) or teacher proxy (TP)	Authors	No. participants (n), age range, mean (SD)	Criterion	Results
Eight item questionnaire (not detailed)	PP	Bacardí-Gascón et al. (2011)	35, NA	Accelerometry	$r_{s:}$= 62 (p < 0.01)
Outdoor recall questions Outdoor playtime time checklist	PP	Burdette et al. (2004)	250, 2.4 - 4.3 y, 3.6 y	Accelerometry	Check list: $r_{s:}$ = 0.33 (p < 0.01) Recall: $r_{s:}$ = 0.20 (p = 0.03)
Questionnaire coded: Frequency of physical activity: 1 'very often' to 3 'not often' Level of physical activity: 1 'very active' to 4 'inactive'	TP	Chen et al. (2002)	24, 3 - 4 y, 3.8 (0.26) y	Actiwatch Calorie counter (steps/day)	Children classified as 'very active' significantly higher TEE per day than those inactive (p < 0.05)
Pre-school Physical Activity Questionnaire (Pre-PAQ)	PP	Dwyer et al. (2011)	67, 3 - 5.9 y, 3.8 (0.74) y	Accelerometry	Mean differences TPA (Sirard) 45.2 min TPA (Reilly) 20.9 min
Structured activity questionnaire: activity index: hours/week by Kriska et al. (1990)	PP	Goran et al. (1997)	101, 5.3 (0.9) y	DLW and indirect calorimetry TEE - REE= AEE	r_p = ns; r values not reported
Physical activity questionnaire developed by author: indoor/outdoor/school based, mode and duration and intensity recorded low/moderate & MVPA (5-9 METs)	PP or TP	Harro et al. (1997)	62, 4 - 8 y, 7 (0.7) y	HR monitoring: MVPA = HR ≥ 140bpm Accelerometry score	Accelerometry r_s = 0.53 (p < 0.0001) HR ≥ 014bpm $r_{s:}$= 0.40 (p < 0.01)

Details of questionnaire	**Parent proxy (PP) or teacher proxy (TP)**	**Authors**	**No. participants (n), age range, mean (SD)**	**Criterion**	**Results**
Netherlands Physical Activity Questionnaire (NPAQ): proxy report: numerical score.	PP or TP	Janz et al. (2005)	204, 4 - 7 y, 5.7(0.5) y	Accelerometry	r_s = 0.33 for TPA; r_s = 0.36 for VPA.
Mother and teacher direct evaluation (record in a questionnaire). 5 category of activity 1) inactive 2) relatively inactive 3) medial 4) relatively active 5) active	PP or TP	Nishikido et al. (1982)	49, 5 - 6 y	Direct observation: AAR- 4 categories. Pedometry step rate	Kendall's rank correlation overall higher for teachers than parents. Teacher evaluation: r = 0.25 with pedometry; r = -0.19 to 0.27 for observed activity. Parents evaluation: r = 0.14 with pedometry; r = -0.14 to 0.12 with observed activity.
Children's Leisure Activities Study Survey (CLASS)	PP	Telford et al. (2004)	58 5 -6 y 5.3 (0.5)	Accelerometry	MPA r_s = -0.06; VPA r_s = -0.04; TPA r_s = -0.04
Seven day diary adapted from Children's Leisure Activities Study Survey (CLASS)	PP	Wen et al. (2010)	31, 3 - 5 y, 3.5 y	Accelerometry	TPA $r_{s:}$= 0.23 (ns); Sed $r_{s:}$ = 0.24 (ns)

AAR: Activity Appearance Rate; AEE: activity energy expenditure; bpm: beats per minute; DLW: doubly labeled water; HR: heart rate; NA: not available; ns: not significant; PP: parent proxy; REE: resting energy expenditure; r_s: Spearman correlations; r_p: Pearson's correlation; TEE: total energy expenditure; TP: teacher proxy; TPA: total physical activity; VPA: vigorous physical activity.

It can be seen that at best there are moderate correlations between the proxy reporting and the secondary methods against which they are validated. However, it has been argued that as the secondary methods often quantify different dimensions of physical activity than what a subjective method tries to quantify, higher correlation values cannot necessarily be expected (Harro 1997).

Despite the limitations of proxy questionnaires, they are useful in situations where objective methods are not available (Burdette et al. 2004) or when used as an addition to an objective method to provide contextual information about physical activity and sedentary behaviours. This information is important in gaining a better understanding of the physical activity and sedentary behaviour habits of pre-school children (Dwyer et al. 2011). In this thesis however, proxy reporting was not adopted as the concern was with objective measurement of physical activity and sedentary behaviour.

1.5.3 Primary methods of measuring physical activity

The primary measures of indirect calorimetry, doubly labeled water (DLW) and direct observation have frequently been used in studies of pre-school children as the criterion method of physical activity. Which approach is adopted is dependent on the elements of physical activity that the researcher wishes to investigate and there are valid arguments to support each approach (McNamara et al. 2010). Direct observation methods are able to provide descriptive information on physical activity behaviour, such as the type of activity, the frequency, duration and intensity of activity, and they are often regarded as being the ‘gold standard’ method for measuring physical activity in young children (Loprinzi and Cardinal 2011; Oliver et al. 2007b). Indirect calorimetry and DLW are both accurate at measuring energy expenditure (Oliver et al. 2007b).

1.5.3.1.1 Indirect calorimetry and doubly labeled water

During indirect calorimetry methods, energy expenditure is measured by the analysis of respiratory gas exchange within a sealed chamber or within a ventilated hood system worn by subjects (Muller and Bosy-Westphal 2003). During the process oxygen consumption (VO_2) and carbon dioxide (CO_2) production are measured and converted to energy expenditure using a formula (Levine 2005; Muller and Bosy-Westphal 2003).

Both indirect calorimetry and DLW techniques require expensive equipment or facilities and are often limited to a laboratory setting. There is some evidence supporting the feasibility of whole room calorimetry with pre-school children (Oortwijn et al. 2009) making it suitable as a criterion measure in the short term. Portable indirect calorimetry equipment is also available for measuring energy expenditure allowing for measurement of free-play. However, the equipment is cumbersome to wear and as a result, this may impact on the child's 'usual' movement patterns and subsequently alter their energy expenditure (Oortwijn et al. 2009). It is therefore more challenging to replicate children's typical activity in their usual environment using these methods.

DLW techniques usually involve participants drinking water which is 'tagged' with a non-radioactive isotope ($D_2{}^{18}O$) (Muller and Bosy-Westphal 2003). Samples of urine, saliva or blood are then collected over 7 - 21 days and changes in concentration of the isotope in body's water allows for calculation of VO_2 consumption and EE (Levine 2005). DLW is argued to be the 'gold standard' for measuring total energy expenditure (TEE) in free-living conditions (Ekelund et al. 2001), and this method has been used successfully with pre-school children to provide a measure of TEE (total energy expenditure) collected over several days (Lopez-Alarcon et al. 2004; Montgomery et al. 2004; Reilly et al. 2004). In the study by Montgomery et al. (2004) the TEE values were used to calculate the child's physical activity level (PAL) (PAL = TEE/predicted resting energy expenditure (pREE)). PAL was found to be positively associated with accelerometry measures of light intensity physical activity and negatively associated with sedentary behaviour (Montgomery et al. 2004).

Limitations of the DLW technique are that it is not able to provide information on the different dimensions of physical activity behaviour such as the intensity, frequency or duration of physical activity undertaken (Rowlands et al. 1997). In addition the use of predicted methods for calculating REE has been shown to have variable accuracy with pre-school children (Finan et al. 1997) and the protocol for directly measuring REE requires fasting in excess of 4 hours and lying still for 30 minutes (Ventham and Reilly 1999) which young children may have difficulty complying with (Montgomery et al. 2004).

The validity of indirect calorimetry and DLW methods to predict energy expenditure has been investigated in a handful of studies since 1993 with pre-school children. Table 1.4 outlines some of the relevant studies undertaken with the pre-school population. Earlier studies of DLW have been reviewed by Goran (1994). In the review, Goran (1994) reports on the finding that DLW has been validated by whole room indirect calorimetry and is accurate within 10% (Goran 1994).

Both indirect calorimetry and DLW methods have been used extensively as criterion methods of energy expenditure in validation studies of secondary measures of physical activity with pre-school children (Lopez-Alarcon et al. 2004; Montgomery et al. 2004; Oortwijn et al. 2009; Pate et al. 2006; Puhl et al. 1990; Reilly et al. 2004). However physical activity is only one variable that contributes to energy expenditure. For example, in a study of 5-year old children, Fontvieille et al. (1993) reported that physical activity accounted for only 16 (7) % (mean (SD)) of TEE over 7 days of monitoring. In their review Oliver et al. (2007b) highlighted the influence that genotype, ethnicity, body weight and obesity have on energy expenditure and that the relationship is not straightforward. Energy expenditure is a complex variable and as already discussed it is a different construct from physical activity. Given the focus of this thesis is on physical activity as opposed to energy expenditure, as well as the restriction to free-living that indirect calorimetry methods may incur and the inability to measure domains of physical activity (with DLW), the DLW and indirect calorimetry methods were not felt to be suitable for use in the current thesis.

Table 1.4: Summary of doubly labeled water and indirect calorimetry studies to predict EE with pre-school children.

Approach	Author	No. participants (n), age range, mean (SD)	Protocol	Results/conclusions
Indirect calorimetry	Finan et al. (1997)	113, 3.9-7.8 y	Indirect calorimetry used to measure REE (mREE) this was compared with predicted REE (pREE) using regression analysis	Most prediction equations do not accurately predict mREE except the FAO/WHO/UNH equation with 99% of predictions being within 200kcal/day
DLW Indirect calorimetry	Goran et al. (1997)	Study 1: 101, 5.3 (0.9) y, Study 2: 68, 6.3 (0.9) y,	Energy expenditure measured with DLW and REE with indirect calorimetry FM & FFM bioelectrical resistance & skin fold	AEE significantly correlated with FM &FFM (r = 0.32) and with weight (r = 0.28 & 0.29 for study 1 & 2); not correlated with FM (r = 0.32)
DLW	Montgomery et al. (2004)	104, 5.4 y, Pre-school age (n): 36, School age (n): 68	Accelerometry data collected over 3 days for pre-schoolers and 7-10 days from school aged. DLW collected d 1 & 7 from pre-schoolers; day 1 & 10 from school aged children. PAL from DLW calculated TEE/pREE	TPA r = 0.33 p < 0.01; LPA: r = 0.31, (p < 0.01); MVPA: r = 0.22, (p < 0.01); Sed r = -0.33, (p < 0.01) with PAL
Whole room calorimetry	Oortwijn et al. (2009)	5, 5.2 (0.4) y	150-175 minute structured protocol within room. EE calculated from measurement of O_2 consumption and CO_2 production by Weirs formula	Feasibility study: conclusion that structured protocol tolerated by young children within a room calorimeter

DLW: doubly labeled water; EE: energy expenditure; FAO/WHO/UNH: Food and Agriculture Organization, World Health Organisation; United Nations University; FM: fat mass; FFM: fat free mass; LPA: light physical activity; mREE: measured resting energy expenditure; PAL: physical activity level; pREE: predicted resting energy expenditure; REE: resting energy expenditure; TPA: total physical activity.

1.5.3.1.2 Direct observation methods

Another primary measure is direct observation which is considered to be a ‘criterion’ measure of physical activity behaviour. While different direct observation methods exist for different settings (physical education classes or free-play) it primarily involves observation of an individual or group of subjects over a period of time by observers who classify subjects’ physical activity behaviours into distinct categories (Vanhees et al. 2005). Direct observation is thought to be particularly suited as a criterion measure of physical activity of young children with whom interpretation of energy expenditure can be difficult (De Bock et al. 2010; Loprinzi and Cardinal 2011; Oliver et al. 2007b).

This method has been argued by some authors to be an objective method (Dollman et al. 2009) and by others to be a semi-objective method (Lubans et al. 2011). Direct observation methods have also been referred to as a subjective method as they rely on humans to observe and record physical activity behaviour (Oliver et al. 2007b). The limitations of direct observation methods are that they are time-consuming with high investigator burden, and therefore these methods are resource intensive (Oliver et al. 2007b). There is also potential for reactive behaviour from participants as a consequence of being observed, meaning that ‘typical’ activity behaviour may not be witnessed. The high investigator burden means that it is not practical for large population-based studies, but it has been used extensively as a criterion method in studies with young children to validate other physical activity measurement methods such as proxy-report (Nishikido et al. 1982), accelerometry (Oliver et al. 2009; Reilly et al. 2003; Sirard et al. 2005; Van Cauwenberghe et al. 2011) and pedometry (Beets et al. 2005; Duncan et al. 2011; Oliver et al. 2007a). There are different objective measurement systems for scoring and recording physical activity and Table 1.5 presents a summary of the direct observation systems which have been used in studies of children.

Table 1.5: Summary of studies investigating the validity and reliability of direct observation methods.

Measurement system	Author	No. participants (n), age range, mean (SD)	Observation Strategy	Criterion Measure	Validity	Reliability
CARS: Children's Activity Rating Scale	Puhl et al. (1990)	Study 1 : 91, 3 - 6 y Study 2: 25, 5 - 6 y	1 min partial time sampling with 5 categories	VO_2 (portable metabolic unit) & HR	Study 2 Significant difference for each level of CARS ($p < 0.05$)	Study 1 84.1%*
CPAF Children's Physical Activity Form	O'Hara et al. (1989)	36, 3rd - 5th grade students	1 min partial time sampling with 4 categories	HR	$r = 0.26$ - 0.90 mean $r = 0.64$	96-98%*
SOCARP System for Observing Children's Activity and Relationships during Play	Ridgers et al. (2010)	Intra observer reliability (n = 14), Inter observer reliability (n = 27), Validity (n = 99) School children	10 s sampling period followed by 10 s recording interval. Level of intensity: 5 categories (same as those used in SOFIT).Type: sport, active games, sedentary activities, locomotion. Group size: alone, small, medium, large. Interaction: prosocial, antisocial. Child observed for 10 min	Accelero metry	$r = 0.67$, ($p < 0.01$)	87-93%
SOFIT System for Observing Fitness Instruction Time	McKenzie et al. (1997)	3rd - 5th grade students	Group observation 10 min momentary time sampling with 5 categories	HR	$r = 0.80$ - 0.91 ($p < 0.01$)	92%*
SOFIT-P System for Observing Fitness Instruction Time for Pre-schoolers	Sharma et al. (2011)	Study 1: 67 Study 2: 27 3 - 6 y	Group observation 10 min momentary time sampling with 5 categories	Accelero metry	Study 2 $r = 0.50$ - 0.54 ($p < 0.01$)	Study 1 > 75%

Measurement system	Author	No. participants (n), age range, mean (SD)	Observation Strategy	Criterion Measure	Validity	Reliability
SOPLAY System for Observing Play and Leisure Activity in Youth	McKenzie et al. (2000)	138, 6^{th} - 8^{th} grade students	2, 3 or 5 min group coding 3 categories plus codes for context, accessibility, supervision & equipment	Accelero metry	*r* = 0.24 - 0.57 (*p* < 0.01)	ICC = 0.74
BEACHES Behaviors of Eating and Activity for Children's Health: Evaluation System	McKenzie et al. (1991)	Study 1: 42, 4 - 8 y, Study 2: 19, 4 - 9 y	1 min momentary time sampling with 5 categories 25 s observation with 35 s for coding	HR	Study 1 HR increased with activity, relationship with HR not calculated	Study 1 94% - 99%* kappa 0.90
FATS Fargo Activity Time-Sampling Survey	Klesges et al. (1984)	14, 2 - 4 y,	3 s continuous time sampling with 30 categories	VO_2 (portable metabolic unit)	*r* = 0.78 - 0.90	91-98%* kappa 0.91
OSRAC-P Observational System for Recording Physical Activity in Children-Preschool version	Brown et al. (2006)	NA	5 s observation followed by 25 s coding. (Two observations per min)	No validation data	No validation data	kappa < 0.80

*HR: heart rate; NA: not available; *Inter-rater percentage agreement.*

Most of the direct observation methods are designed to assess an individual participant's physical activity behaviour. However, the System for Observing Play and Leisure Activity in Youth (SOPLAY) has been developed for the purpose of assessing group levels of physical activity (Saint-Maurice et al. 2011). Saint-Maurice et al. (2011) investigated the concurrent validity of the SOPLAY in a study of 160 children aged 9 - 12 years. The authors concluded that while there was support for the validity of the SOPLAY care should be taken in interpreting the intensity level of activity, in particular with category 2 activities which are described as 'walking'. Although walking can be characterised as either a light or a moderate intensity activity, Saint-Maurice et al. (2011) suggest that interpreting walking as MVPA would lead to an overestimation of time that children spend in MVPA. The authors argue that walking was more typically a light intensity physical activity (Saint-Maurice et al. 2011). This is important, as in other direct observation scales, walking activities have been split into different categories depending on the speed of walking undertaken. For example in the CARS there is 'slow' and 'brisk' walking (Puhl et al. 1990). The interpretation of whether a walking activity is categorised as a 'light' or a 'moderate' intensity activity is not consistent between studies.

Of the direct observation methods available the Children's Activity Rating Scale (CARS) and the Children's Physical Activity Form (CPAF) have been most commonly used in studies of young children (DuRant et al. 1994; Finn and Specker 2000; Kelly et al. 2004; Oliver et al. 2007a; Oliver et al. 2009; Reilly et al. 2003; Van Cauwenberghe et al. 2011). Unlike the CPAF, which does not have established reliability (Sharma et al. 2011) and has concurrent validity against heart rate, a secondary measure of physical activity, in older children (8 - 10 y), the CARS has been shown to have good inter-rater reliability with younger children (aged 3 - 4 y) (DuRant et al. 1993). The CARS has also been validated against indirect calorimetry for use in a free-play conditions with younger children aged 5 to 6 years (Puhl et al. 1990).

The CARS is limited by the fact that it does not provide any information on the context of activity (Sharma et al. 2011). However, the 'Observational System for Recording Physical Activity in Children-Pre-school version' (OSRAC-P), developed by Brown et al. (2006) and modelled on the CARS, collects additional data on the topography of activity (e.g. running, walking, sitting), as well as the social circumstances (i.e. initiates activity, prompt for activity peer or teacher), and non-social environmental circumstances (inside, outside, transition) of children's activity. The addition of this information would be useful for descriptive purposes, particularly for research into the impact of physical activity interventions. Although the authors reported on good inter-observer agreement for all categories (kappa > 0.80) except group (kappa = 0.79) using the kappa statistic, there was a wide variation in the observer agreement for physical activity and type of activity (kappa = 0.18 - 1.00 and 0.50 - 1.00 respectively). This variation highlights the vulnerability of direct observation methods as they reply on humans to accurately observe, interpret and record data (Oliver et al. 2007b). The collection of this additional information may in part account for lower reliability scores seen in the study by Brown et al. (2006), in contrast to the study by Puhl et al. (1990), where inter-rater reliability was 84.1 (10)% (mean percentage agreement (SD)). Interestingly, most of the observer disagreements occurred for Level 1 and Level 2 activities, which are the stationary activities either with arm movements (Level 2) or without arm movements (Level 1). The authors suggest combining these into a composite ‘sedentary measure’ for future analysis, arguing that there is limited need for a differentiation between these categories (Brown et al. 2006).

The OSRAC-P scale has yet to be validated and it depends on the purpose of the study whether the descriptive information on the typography and context of physical activity is desired. Coding the additional information in the OSRAC-P may detract from accurate coding of the observed physical activity behaviour, with consequences for the reliability of the tool.

The Children’s Activity Scale (CARS), developed by Puhl et al. (1990), allows minute-by-minute coding of children’s activity into five different categories of intensity level (Table 1.6). In the study by Puhl et al. (1990) (n = 25, 12 boys, 13 girls, 5 - 6 y), they were able to discriminate between the categories of activity levels against energy expenditure measured by indirect calorimetry (VO_2) and heart rate.

Table 1.6: Children’s Activity Rating Scale (CARS).

Level	Description
1	Stationary/non-moving e.g. sitting quietly
2	Stationary/with movement e.g. sitting/standing swinging arms, standing still
3	Translocation slow/easy e.g. slow walk
4	Translocation medium/moderate e.g. moderate walk
5	Translocation fast/very fast/strenuous e.g. running

(Danner et al. 1991; Puhl et al. 1990)

Levels four and five represent moderate-to-vigorous intensity physical activity (activities with energy cost at least three times resting energy expenditure) (Sirard et al. 2005).

A limitation of the CARS is that as it records activity once during a 1-minute period, and then averages the scores to obtain a mean score, it may not accurately reflect the activity undertaken during that 1-minute period. Oliver et al. (2007a) uses the example of a child being engaged in sedentary behaviour for 3 s (Level 1 activity), then engaging in 57 s of vigorous activity (Level 5 activity). This would result in categorising the minute at Level 3 ((1+5)/2 = 3), which would mean the minute is classified as 'light' intensity physical activity, and thus misrepresenting the child's true activity. However, given the intermittent nature of physical activity reported in young children (Bailey et al. 1995) it is possibly unlikely that they would sustain vigorous activity for 57 s.

Sirard et al. (2005) adapted the CARS to a 15-s sample period and applied it in a study of pre-school children. They combined Levels 1 and 2, which are both stationary activities, into one 'sedentary' category. The researchers reported high ICC values for intra-observer agreement (0.95 to 0.96 and 0.88 to 0.94 for the start and end of data collection period) and a high percentage agreement (75 to 99%) of 15 s physical activity categorisation across all time points (kappa = 0.66 to 0.98) (Sirard et al. 2005). The shorter sample period may go some way to reduce the misrepresentation of activity for that observed period. Oliver et al. (2009) have gone on to use a second-by-second CARS coding system, where codes are applied for each second and then an average is calculated for a 15-s period. While they reported good intra-observer reliability (96%) with this approach there is the possibility that coding for each second could increase the risk of introducing operator bias or error.

In summary, while the CARS is not suitable for large scale studies due to high operator burden, it is considered to be a criterion method of choice for use as the 'gold standard' method in studies of physical activity in young children (Loprinzi and Cardinal 2011). In this thesis, the CARS was selected as the criterion measure of physical activity against which accelerometers were validated.

1.5.4 Secondary methods of measuring physical activity

Secondary methods of measuring physical activity include heart rate and activity monitors such as pedometers and accelerometers. Each will be discussed in this section.

1.5.4.1.1 Heart rate

Heart rate (HR) is commonly used as an objective measure of physical activity (Eston et al. 1998). However, Oliver et al. (2007b) argues that HR is particularly limited as being a 'proxy' measure of physical activity as it is only an indirect approximation of energy expenditure. Accurate measurement of EE from HR relies on the linear relationship between oxygen uptake (VO_2) and HR (Melanson and Freedson 1996).

There are several limitations with using heart rate as a measure of physical activity. Firstly, HR responses following movement are delayed, which may mask the intermittent behaviour of physical activity (Trost 2001) and this may be important for studies of young children. The relationship between HR with high and low intensity physical activity is weak (Loprinzi and Cardinal 2011) and it does not accurately measure sedentary behaviour (De Bock et al. 2010). Another key limitation with using HR is that it may be affected by factors other than physical activity, such as emotions, surroundings, temperature and dehydration (Muller and Bosy-Westphal 2003) as well as cardiorespiratory fitness (Trost 2001).

To overcome some of the limitations with HR, techniques have been developed including the use of relative HR indices which adjust for differences in fitness and age (Trost 2001). The most commonly used indices are the Physical Activity Heart Rate-25 (PAHR-25) and the Physical Activity Heart Rate-50 (PAHR-50) which relate to the percentage of time spent with HR at 1.25 (light to moderate) and 1.50 (vigorous) times resting heart rate (RHR) (Logan et al. 2000). However, the HR indices rely on accurate measurement of RHR and the problem is that several different definitions of RHR have been published (Trost 2001). In a study of 20 children (10 boys, 10 girls, mean (SD) age: 4.4 (0.4) y), Logan et al. (2000) highlighted the problem of using different definitions of RHR when they compared measured RHR with derived RHR using four commonly used definitions of RHR. The derived RHR values were then converted to the PAHR-25, PAHR-50 and average activity heart rate (AHR) indices. As well as there being significant differences in the derived estimates of RHR depending on the definitions of RHR used, the PAHR-25 varied by 10 -50%, the PAHR-50 by 16 - 65% and AHR varied by 9 - 44% depending on the derived definition of RHR used. Therefore, in order to use these indices in a meaningful way, a consensus on how RHR is defined and measured first needs to be reached (Logan et al. 2000).

In summary, if estimation of energy expenditure is of interest then HR may be of value and there are some promising results from studies with pre-school children where HR has been used in combination with accelerometry to measure energy expenditure (De Bock et al. 2010; Ojiambo et al. 2012). There are many advantages to using HR in free-living conditions as the instruments used to record HR are low-cost, easy-to-use, unobtrusive, and they have the capability of recording data over an extended period of time (Loprinzi and Cardinal 2011). In addition, in one of the few studies to have used HR with young children it has been found to be a reliable measure of physical activity (DuRant et al. 1992). However, as the relationship between VO_2 and HR is individual for intensity, there is a need for individual calibration which can be time consuming and would make HR impractical for use in large population-based studies. Finally, as the focus of this thesis is on physical activity behaviour as opposed to energy expenditure, HR monitoring was not included as a measure of physical activity.

1.5.4.1.2 Pedometers

Pedometers are one type of activity monitor which are inexpensive, easy-to-use and provide objective physical activity data on the frequency of activity in a user-friendly way (e.g. number of ‘steps’ taken per day) (Tudor-Locke et al. 2004). Pedometers are small, lightweight motion sensors which measure movement of the body in the vertical plane (Pate et al. 2010) and are usually worn on the hip. These features of a pedometer have meant that they are well suited for data collection from large populations. However, while pedometers are very useful in providing data on a person’s ambulatory activity level (Crouter et al. 2003), they are not sensitive to gait differences such as stride length which vary from person to person (Zhang et al. 2003). They are also unable to take account of the additional energy expenditure from activities such as uphill walking, stair climbing and upper limb activity (Bassett et al. 2000). Importantly, pedometers record data in cumulative steps, and as such do not provide information on patterns of physical activity, such as the intensity, frequency and duration of physical activity, and they do not have storage capabilities (de Vries et al. 2009).

There are several models of pedometer available, however, the Digi-walker (SW-series; Yamax Co., Yasama Corp, Tokyo, Japan) is the only model to have been used in validation studies with pre-school children. Table 1.7 presents a summary of the validation studies of pedometers undertaken with pre-school children. Unfortunately, information on the reliability of the Digi-walker was not available in the studies with pre-school children.

However in a review, de Vries (2009) reported that pedometry studies with older children suggest high intra-instrument and inter-instrument reliability with the Digi-walker pedometer.

The criterion validity of the Digi-walker has been reported in four studies and construct validity in one study (Table 1.7). The correlation results suggest moderate associations with direct observation and good construct validity with accelerometry. This supports the findings of de Vries (2009) for the pre-school age group, where they suggest that there is a modest level of evidence for the validity of pedometers in the pre-school age group.

However, Oliver et al. (2007b) highlights the limitation of correlation analysis for assessing validity, which measures the strength of a relationship but can overlook systematic differences. In their small study of 13 pre-school children (mean (SD) age: 4.1(0.6) y) they used the Bland and Altman method to compare step counts when walking at a slow speed, a normal walk speed and run speed, from pedometers worn on each hip and the back, against the CARS direct observation scale (Oliver et al. 2007a). The authors reported wide limits of agreements (15 - 44 steps over 29 m) for all ambulatory conditions and pedometry positions, leading them to question the accuracy of pedometers as a measure of physical activity in pre-school children. However, these findings were based on a small sample (n = 11) and make use of the CARS coded at 60 seconds, which the authors concluded could limit the accuracy of the CARS to act as a criterion measure of physical activity (Oliver et al. 2007a).

In summary, while pedometers may be useful for a general assessment of accumulated ambulatory activity (number of steps taken), they are unable to provide detail on the dimensions of physical activity e.g. intensity, frequency and duration of physical activity. The inability of pedometers to store information over extended periods of time also limits their use for assessing habitual physical activity behaviour. With this and the moderate evidence of the validity of pedometers for use with pre-school children, they were not selected as the instrument to measure habitual physical activity behaviour of pre-school children in the current thesis.

Table 1.7: Summary of validation studies of pedometry with pre-school children.

Pedometer model	Study authors	No. participants (n), age range	Method	Criterion	Validity
Digi-walker	Cardon and Bourdeaudhuij (2007)	76, 4 - 5.9 y	4 days wearing pedometer and accelerometer, correlation of steps count with MVPA min.	Accelerometry	*r* = 0.73
Digi-walker	Louie and Chan (2003)	148, 3 - 5 y	Observed during 25 min free-play	CARS	*r* = 0.64
Digi-walker	McKee et al. (2005)	30, 3 - 4 y	Observed during 1 hour of self-selected activities	CARS	*r* = 0.64 to 0.95
Digi-walker	Nishikido et al. (1982)	49, 5 - 6 y	Direct observation over 2 days in 2 nurseries AAR categories: sit, stand, walk, run	AAR	*r* = -0.50 & -0.77 sitting *r* = -0.37 & -0.21 standing *r* = 0.36 & 0.42 walking *r* = 0.69 & 0.83 running
Digi-walker	Oliver et al. (2007a)	13, 3 - 5 y	Observed during 35 min free-play	CARS	*r* = 0.54, (*p* = 0.04) LOA (15 - 44 steps over 29 metres)

AAR: Activity Appearance Rate; CARS: Children's Activity Rating Scale; LOA: limits of agreement; MVPA: moderate-to-vigorous activity.

1.5.4.1.3 Accelerometers

Accelerometers are another type of activity monitor which can be used to collect objective data on physical activity behaviour. Recent technological advances have seen a dramatic increase in the use of accelerometers within physical activity research (de Vries et al. 2009; Rowlands 2007).

Accelerometers measure acceleration of the body part to which they are attached (de Vries et al. 2009). The sensors in early accelerometer models were a piezoelectric element and a seismic mass attached to a piezoceramic cantilever arm (Chen and Bassett 2005). More recent generations of accelerometers contain a Micro-Electro-Mechanical System (MEMS) solid state accelerometer (John et al. 2010). Both sensors detect acceleration or change in velocity over time (Corder et al. 2008). The theoretical basis for their use in physical activity measurement is that acceleration of the body is directly proportional to the muscular forces generated, which relates directly to energy expenditure (Freedson and Miller 2000), although their ability to accurately estimate energy expenditure in young children is still questionable (de Vries et al. 2006).

Several reviews have concluded that accelerometers provide a reliable, valid and practical means of objectively measuring physical activity (Corder et al. 2008; Oliver et al. 2007b; Reilly et al. 2008). Accelerometers have the advantage over HR and pedometry in being able to provide information on the frequency, intensity and duration of physical activity undertaken. Recent technical advances in memory capacity and battery life means that data can be stored over extended periods of time, with up to 356 days of data collection possible with some models (de Vries et al. 2009). Unlike direct observation methods, there is low researcher burden with accelerometers and they are free from researcher bias (Oliver et al. 2007b). Compared to methods such as DLW and indirect calorimetry, accelerometers are relatively inexpensive, light weight and unobtrusive (usually worn clipped to a waist-band) and are reported to be tolerated by children, with low reactivity (de Vries et al. 2006; Vincent and Pangrazi 2002). However, similar to pedometers, accelerometers are unable to take into account the increased energy expenditure with upper limb movements (unless attached to the arms) or to assess activities such as cycling or walking on a graded terrain (Janz 2006). Despite these limitations accelerometry is the most frequently used method for objectively measuring physical activity in research with pre-school children (Oliver et al. 2007b).

Accelerometers can be used to measure acceleration of the body in one, two or three orthogonal planes (vertical, mediolateral and anteroposterior planes). Several models of accelerometer are commercially available, however, the most frequently used accelerometer to date in studies of pre-school children has been the uniaxial Actigraph accelerometer, either the 7164 Actigraph model (Actigraph, Fort Walton Beach, FL, USA) or the GT1M model (Actigraph, Fort Walton Beach, FL, USA) (de Vries et al. 2009; Pate et al. 2010; Trost et al. 2011). The 7164 Actigraph model was earlier known as the Computer Science Applications 7164 model (CSA 7164 model) and the MTI 7164 model, and for the purposes of this thesis it will be referred to as the 7164 model.

Uniaxial models, such as the GT1M and the 7164 are usually worn on a belt around the waist in a position which aligns the sensitive axis with the vertical plane. The *activ*PAL™ (PAL technologies Ltd, Glasgow, UK) is an activity logger which incorporates a uniaxial accelerometer and is attached to the thigh mid-way between the hip and knee (Godfrey et al. 2007; Godfrey et al. 2008). The advantage of this positioning is that the accelerometer is able to distinguish between static postures for example sitting/lying, and standing, and the dynamic activity of movement (Godfrey et al. 2007).

Omnidirectional accelerometers are available, such as the Actical (Mini Mitter Co., Inc., Bend OR, USA) and the Actiwatch (Mini Mitter Co., Inc., Bend OR,USA). Omnidirectional accelerometers are sensitive to motion in any direction but are most sensitive to acceleration in the vertical plane (Chen and Bassett 2005). It has been argued that as omnidirectional accelerometers can only measure acceleration in one plane at a time, that they therefore function as a single axis accelerometer (Pate et al. 2010; Pfeiffer et al. 2006). Also available are triaxial accelerometers, such as the RT3 (Stayhealthy, Inc., Monrovia, CA, USA), Tritrac-R3D (Hemokinetics, Inc./Professional Products, Division of Reining International Ltd., Madison, WI, USA) and the GT3X and GT3X+ (Actigraph, Fort Walton Beach, FL, USA) which measure acceleration in three planes (vertical, AP, ML). These accelerometers provide a value for acceleration recorded in each plane, as well as a composite value known as the vector magnitude value (de Vries et al. 2009).

Although the more recent GT3X model from Actigraph has the capacity to provide raw acceleration data in a filtered or unfiltered mode, in many studies the data output are expressed in terms of activity counts which are collected over an epoch period (Freedson et al. 2012). The epoch is set prior to data collection and is usually between 1 and 60 seconds.

Accelerometry counts are arbitrary, having no biological meaning per se (Reilly et al. 2008) and even when subjected to the same acceleration the count outputs from different types of accelerometers are not comparable due to the differences in the transducers, sampling frequencies and signal filters of different accelerometer models (Chen and Bassett 2005; Chen et al. 2012). Calibration studies have therefore been undertaken to convert the count output into a biologically meaningful format (Freedson et al. 2005). This has resulted in establishing ranges of counts (cut-points) to correspond with different intensity thresholds (Freedson et al. 2005) allowing researchers to express count data as time spent at different intensities (Bassett et al. 2012). However, a wide range of varying cut-points has been proposed and there is no consensus on which is appropriate (Ekelund et al. 2011; Reilly et al. 2008). Some accelerometer models also provide data on energy expenditure, however it is argued that these data are not accurate for children under the age of 10 years, as the calculations are based on adult values of energy expenditure (de Vries et al. 2006). Finally, some researchers have used accelerometry count data to predict energy expenditure using prediction equations. While these prediction equations seem valid for physical activity undertaken by children in a laboratory-based setting, they are not necessarily valid for use in free-living conditions (Kristensen et al. 2008; Nilsson et al. 2008). Therefore, time spent at different intensities using the accelerometry counts is most frequently reported in the literature. How the count data should be interpreted, however, is a contentious issue and the specific concerns associated with data processing, analysis and interpretation is one area which this thesis seeks to address (Cliff et al. 2009b).

Despite the limitations with accelerometers, they have been used and continue to be used to measure habitual physical activity of children in several large-scale population-based studies such the Millennium Cohort Study in the UK (Basterfield et al. 2011) and the National Health and Nutritional Examination Survey (NHANES) in the USA (Troiano et al. 2008).

Table 1.8 Summarises studies which have investigated the validity and reliability of accelerometers with pre-school children.

Table 1.8: Summary of validation studies of accelerometers with pre-school children.

Accelerometer make & model	Author	No. participants (n), age range, mean (SD)	Method	Criterion method	Validity	Reliability
Actigraph (GT1M)	De Decker et al. (2013)	52, 4 - 6 y	5 day free-living 1 hour classroom observation	Direct observation	Sed: Sen: 58.5%; Sp: 61.2% ROC:0.6; Sed including standing: Sen: 46.3%; Sp: 75.8%	NA
Actigraph (7164)	Fairweather et al. (1999)	11, 3.7 (0.5) y	40 - 50 min of structured play.	CPAF	$r = 0.87$ ($p < 0.01$)	$r = 0.98\text{-}0.99$
Actigraph (7164)	Hands et al. (2006)	24, 5 - 6 y	30 min free-play over 5 days	CARS	$r = 0.5$ (n = 24) $r = 0.8$ (n = 23) when child on swing removed from analysis).	NA
Actigraph (7164)	Kelly et al. (2004)	78, 3 - 4 y	35 - 45 min of structured play.	CPAF	$r = 0.72$ ($p < 0.001$)	NA
Actigraph (7164)	McIver et al. (2005)	30, 4.4 y	Structured activity and free-play. Worn on hip and back.	VO_2 (portable metabolic unit)	Rest and structured: Hip position $r = 0.78$ ($p = 0.018$); back position $r = 0.68$ Free-play: hip position $r = 0.35$; back position $r = 0.45$	Inter-instrument reliability $r = 0.73$ & $r = 0.78$
Actigraph (7164)	Montgomery et al. (2004)	104, 2 - 7 y	3 days with pre-school children, 7 - 10 days with school aged children % time in LPA; % time in MVPA;% time in Sed	DLW	Sed: $r = -0.33$ LPA: $r = 0.31$ MVPA: $r = 0.22$ TPA: $r = 0.33$	NA
Actigraph (7164)	Pate et al. (2006)	29, 3 - 5 y, 4.4 (0.8)y	Rest and structured activities, cross validation with unstructured activities indoor and outdoor	VO_2 (portable metabolic unit)	VO_2 and counts $r_s = 0.82$	NA

Accelerometer make & model	Author	No. participants (n), age range, mean (SD)	Method	Criterion method	Validity	Reliability
Actigraph (7164)	Reilly et al. (2003)	50, 3 - 4 y	In nursery Free-play activities.	CPAF	Sed: Sen: 83 (14)%; Sp: 82 (11)%	NA
Actigraph (7164)	Reilly et al. (2006a)	85, 4.6 (1.1) y	3 - 7 days Free living	DLW	Mean error (LOA) with Ekelund et al. equation: +0.3 $MJ{\cdot}d^{-1}$ (-3.7 to 4.3) Puyau et al. equation: -0.3 $MJ{\cdot}d^{-1}$ (3.2 to -3.8)	NA
Actigraph (7164)	Sirard et al. (2005)	269, 3 - 5 y, Reliability: 16, 3 - 5 y	1 hour over 1 - 3 days in nursery. Free-play	CARS	$r_p = 0.46$ to 0.70 ($p < 0.001$)	ICC = 0.84
Actigraph (7164)	Toschke et al. (2007)	11, 5 - 6 y	5 days free-living. Two accelerometers: one over hip & one over umbilicus	NA	NA	Inter instrument reliability r = 0.95 however umbilicus were 1.5 (+50cpm) higher
Actiheart	Adolph et al. (2012)	64, 4.5 (0.8) y	3 hour of play, slow, moderate and fast translocation	VO_2 (whole room calorimetry)	True positives predictive rates: Sed: 75% LPA: 61% MVPA: 82%	NA
Actical	Adolph et al. (2012)	64, 4.5 (0.8) y	3 hour of play, slow, moderate and fast translocation	VO_2 (whole room calorimetry)	True positives predictive rates: Sed: 77% LPA: 63% MVPA: 69%	NA

Accelerometer make & model	**Author**	**No. participants (n), age range, mean (SD)**	**Method**	**Criterion method**	**Validity**	**Reliability**
Actical	Evenson et al. (2008)	33, 5 - 9 y	Structured play, walk, run.	VO_2 (portable metabolic unit)	Sed: Sen: 97%; Sp: 98% MPA: Sen: 78%; Sp: 79% VPA: Sen: 77%; Sp: 79%	NA
Actical	Pfeffier et al. (2005)	18, 4.4 (0.7) y	Resting, structured activities. 20 min of unstructured activities indoor and outdoors	VO_2 (portable metabolic unit)	VO_2 and counts $r_p = 0.75$	ICC: 0.5
Actical	Pfeiffer et al. (2006)	18, 3 - 5 y, 4.4 (0.7) y	Rest, structured. 20 min of unstructured activities indoor and outdoors	VO_2 (portable metabolic unit)	VO_2 and counts $r_p = 0.89$	NA
*activ*PAL™	Davies et al. (2012a)	Validation study 30, 4.1 y, Reliability study 20, 4.4 y	Study 1: 1 hour direct observation Study 2: 7 consecutive 24 hour period	Study 1: Direct observation	Median values: Sit/lie: Sen: 92.8%; Sp: 97.3% Stand: Sen: 91.8%; Sp: 86.5% Walk: Sen: 77.9%; Sp: 96.5%	ICC ranged from 0.37 to 0.93 With > 0.8 for 5 or more days of monitoring
*activ*PAL™	De Decker et al. (2013)	52, 4 - 6 y	5 day free-living 1 hour classroom observation	Direct observation	*activ*PAL™: Sed: Sen: 53.8%; Sp: 67.5%; ROC: 0.61 Sed including standing: Sen: 27.8%; Sp: 75.8%; ROC: 0.52	NA
Actiwatch	Finn and Specker (2000)	40, 3 - 4 y	5 - 6 hour of direct observation free-living	CARS	$r = 0.03 - 0.92$ (mean 0.74) ($p < 0.01$)	NA
Actiwatch	Kelly et al. (2004)	78, 3 - 4 y	35 - 45 min of structured play	CPAF	$r = 0.16$ ($p > 0.05$)	NA

Accelerometer make & model	Author	No. participants (n), age range, mean (SD)	Method	Criterion method	Validity	Reliability
RT3	Adolph et al. (2012)	64, 4.5 (0.8) y	3 hours of play, slow, moderate and fast translocation	VO_2 (whole room calorimetry)	True positives predictive rates: Sed: 76% LPA: 65% MVPA: 79%	NA
Tracmor	Hoos et al. (2003)	11, 3 - 12 y	2 week data collection, free-living activities	DLW	$r = 0.79$ ($p < 0.01$)	NA

CPAF: Children's Physical Activity Form. CARS: Children's Activity Rating Scale; DLW: Doubly labeled water; ICC: inter class correlation coefficient; LPA: light physical activity; MPA: moderate physical activity; MVPA: moderate-to-vigorous physical activity; NA: not assessed; ROC: Receiver Operating Characteristic Curve; r_p*: Pearson's correlation;* r_s*: Spearman's correlation; Sed: sedentary behaviour; Sen: sensitivity; Sp: specificity; VPA: vigorous physical activity*.

Seventeen studies were reviewed. Of these ten investigated the criterion validity of the Actigraph accelerometer. The Actigraph was validated against DLW in two studies (Montgomery et al. 2004; Reilly et al. 2006a), indirect calorimetry in a further two studies (McIver et al. 2005; Pate et al. 2006) and six studies used direct observation as a criterion measure (De Decker et al. 2013; Fairweather et al. 1999; Hands et al. 2006; Kelly et al. 2004; Reilly et al. 2003; Sirard et al. 2005). The results suggest a poor correlation between the Actigraph output and EE as measured by DLW (r = 0.22 - 0.33) (Montgomery et al. 2004), with one study concluding that the Actigraph is inadequate at estimating free-living TEE (Reilly et al. 2006a). There were moderate to high correlations with indirect calorimetry with studies of structured activities (r = 0.68 - 0.82) (McIver et al. 2005). Slightly lower correlation values for the Actigraph and direct observation were reported for free-play activities in one study (r = 0.35 - 0.45) (McIver et al. 2005) but a higher value was reported in another study which included both structured and free-play activities (Pate et al. 2006). In the studies which set out to validate the Actigraph with direct observation, correlations were generally moderate to high (r = 0.46 - 0.87) (Fairweather et al. 1999; Kelly et al. 2004; Sirard et al. 2005), except in one study where the correlation was lower (r = 0.5) (Hands et al. 2006) until one participant who was on a swing was excluded (r = 0.8).

One study investigated the validity of the Actigraph when distinguishing between different sedentary postures (sitting and standing) (Davies et al. 2012a). The results suggested that the receiver operator characteristic (ROC) was poor (0.6), with low sensitivity (58.5%) and specificity (61.2%). The inability to distinguish between postures such as lying, sitting and standing is a known limitation of many models of accelerometers. Being able to distinguish between postures is important if researchers wish to distinguish between sitting and standing as sedentary behaviours. The *activ*PAL™, which is positioned on the anterior aspect of the upper thigh, is able to detect postures as a result of the inclination of the thigh (De Decker et al. 2013) and the *activ*PAL™ has demonstrated good validity with adults (Godfrey et al. 2007; Grant et al. 2006). There are a few validations studies which have used the *activ*PAL™ with pre-school children and in these there are conflicting findings. Davies et al. (2012a) reported good sensitivity and specificity when distinguishing between postures (Sensitivity: 78 - 93%; Specificity: 87 - 97%), while De Decker et al. (2013) reported poor sensitivity and specificity when distinguishing between sitting and standing (Sensitivity 53.8% and Specificity 67.5%). Interestingly, in the same study, the sensitivity and specificity values for the *activ*PAL™ were similar to those reported for the Actigraph. The convergent validity of the *activ*PAL™ and Actigraph evaluated from 7 days of free-living in pre-school children (n = 23, 4.5 (0.7) y) reported small but significant differences between the monitors for sedentary behaviour, but stated that these were acceptable at a group level (Martin et al. 2011). Another possible limitation of the *activ*PAL™ is whether wearing the monitor taped to the front of the thigh would be tolerated by young children. In their study of 52 pre-school children De Decker et al. (2013) remarked that 38% of parents reported their children had skin irritation due to taping the *activ*PAL™ to the thigh.

The validation studies of the other models of accelerometer suggest that the Actiwatch had variable validity against direct observation (r = 0.16 - 0.92) (Finn and Specker 2000; Kelly et al. 2004); the Tracmor had good validity with DLW (Hoos et al. 2003) and validity of the Actical was high against VO_2 (r = 0.75 - 0.89) (Evenson et al. 2008; Pfeiffer et al. 2006).

In the studies which investigated reliability of accelerometers these values were generally high (r = 0.73 to 0.95) for the Actigraph (McIver et al. 2005; Sirard et al. 2005; Toschke et al. 2007). However, in the Toschke et al. (2007) study, while the correlation between two accelerometers was high the monitor positioned at the umbilicus recorded 1.5 times (+50 cpm) higher than a monitor positioned over the hip in data collected from pre-school children over 5 days (n = 11). This highlights the limitation of relying on correlational analysis which can present a strong association which can overlook systematic bias (Oliver et al. 2007b).

The reliability of the Actiwatch was variable (r = 0.37 - 0.93) (Finn and Specker 2000; Kelly et al. 2004) and the Actical had reasonable reliability (ICC= 0.5) (Pfeiffer et al. 2006).

In summary, the Actigraph accelerometer has undergone the greatest number of validation studies with pre-school children in comparison to other models. The results of the validation studies suggest that it has good validity against indirect calorimetry (McIver et al. 2005; Pate et al. 2006) and direct observation measures (Fairweather et al. 1999; Kelly et al. 2004; Sirard et al. 2005). The validity of Actigraph accelerometers to estimate TEE, however, is poor (Reilly et al. 2006a) and therefore the Actigraph is possibly most useful as a behavioural measure of physical activity. The Actigraph has been found to be a reliable instrument for measurement of physical activity in pre-school children, although its ability to distinguish between sedentary behaviours such as standing and sitting is poor. While there are promising results from the *activ*PAL™ to distinguish between different postures (lying/sitting, standing and walking) (Davies et al. 2012a), it is unclear whether the application of the monitor to the thigh might impact on compliance of young children over extended periods of time. It is not clear if the *activ*PAL™ accelerometer performs better than the Actigraph in distinguishing between sedentary and non-sedentary activities in pre-school children (De Decker et al. 2013). Therefore, for the purposes of this thesis the Actigraph accelerometer was selected as accelerometer for measurement of habitual physical activity in pre-school children.

1.5.5 Summary

The decision of which method to adopt for measuring physical activity in young children depends on the purpose and scope of the study. If the interest is in energy expenditure then methods such as DLW and indirect calorimetry are the criterion measures (Oliver et al. 2007b). These methods do, however, require expensive equipment, and indirect calorimetry may not be tolerated or feasible over an extended period of time by young children, limiting their usability. In large populations subjective measures are cheap and easy to use, but are limited for use in young children, as accurate recall is problematic (Pate et al. 2010) and proxy reporting may be subject to reporting bias (Noland et al. 1990). Direct observation is considered to be a criterion method for measuring physical activity in young children (Kelly et al. 2004), however the high investigator burden limits its use for data collection over extended periods of time. Heart-rate monitors, pedometers and accelerometers offer the opportunity for objective measurement of physical activity in large populations in a cost-effective manner. Each method has advantages and disadvantages. However, accelerometry can provide an objective means of measuring of habitual activity over several days which does not require the individual calibration needed for HR monitoring. In addition, accelerometers can collect and store information on the intensity, duration and frequency of physical activity over a period of days or months, a facility which is not available with pedometers. Of the accelerometers available, the Actigraph offers a valid and reliable means of objectively measuring physical activity. There are however, methodological decisions around the use of accelerometers which need to be clarified, as variability in approaches to data collection and analysis can lead to marked differences in estimations of physical activity and sedentary behaviour. The methodological questions concerning the use of accelerometers will be considered in the next section

1.6 METHODOLOGICAL CONSIDERATIONS IN THE USE OF ACCELEROMETERS

1.6.1 Introduction

This final section of the literature review will discuss the outstanding methodological questions concerning the use of accelerometers which this thesis seeks to address.

1.6.2 Which epoch is most accurate?

The accelerometry output of counts is the summation of the filtered acceleration signals over a user-specified time interval which is commonly referred to as an epoch and epochs usually range from 1 s to 60 s (Ayabe et al. 2013; Trost et al. 2005). At the end of each epoch the counts are then stored in the memory of the accelerometer for the duration of the data collection period. The limited memory capabilities of early generation accelerometers meant that the 1 minute sampling period was the shortest epoch which would allow for data to be collected over a number of days (Gabriel et al. 2010). Therefore, prior to 2005 most of the accelerometry studies which explored habitual physical activity of children collected data in 1-minute epochs (Fisher et al. 2005a; Heelan and Eisenmann 2006; Janz et al. 2004; Janz et al. 2010; Kelly et al. 2006; Kelly et al. 2007; Roemmich et al. 2006; Telford et al. 2005). Technical advances have resulted in improved data storage capabilities in modern accelerometers which permits the use of shorter epochs to collect data over extended periods of time (Gabriel et al. 2010).

The implications of using different epochs to collect data on habitual physical activity estimates have been debated within the literature (Cliff et al. 2009b) and the issue is often referred to as an 'epoch effect'. It has been argued that shorter epochs may be particularly important for capturing physical activity behaviour of young children, who are believed to typically engage in short, sporadic bursts of high intensity physical activity, interspersed with varying periods of low or moderate duration activity (Bailey et al. 1995). It is possible that using 1-minute epochs could result in a 'smoothing effect', whereby vigorous activity is obscured and underestimated as a result of 'averaging' a short episode of high intensity physical activity with a longer period of low intensity physical activity in the same epoch (Trost 2001). However, while it has been hypothesised that there may be benefits in using shorter epochs to improve accuracy of estimates of time spent at various intensities of physical activity (McClain et al. 2008), there is no empirical evidence of this for pre-school children and a consensus over which epoch to use has not been agreed (Cliff et al. 2009b).

Studies of school-aged children have investigated the implications of using different epochs for data collection. McGrath and Hinckson (2009) compared Actical accelerometer data collected in 15- and 60-s sampling periods in a study of 79 children (mean (SD) age: 9.7 (0.4) y) over 7 days. The 60-s epoch led to under-recording of vigorous activity by 50% and sedentary behaviour by 20%. Similarly, Nilsson et al. (2002) found a shorter epoch resulted in significantly more time in high (p < 0.01) and very high (p < 0.01) intensity activities in data collected in 5-s epochs re-integrated into 10-, 20-, 40- and 60-s epochs from children (n = 16, 7 y), measured over a 4-day period using a 7164 accelerometer. However, there was no significant epoch effect for time spent at moderate intensity between any of the epoch settings.

Rowlands et al. (2006), also found that the number of minutes spent in very vigorous activity was significantly underestimated when 60-s epochs were used as opposed to 1-s epochs with the RT3 accelerometer. Although using 1-s epochs may be advantageous, the RT3 can only collect vector magnitude data (a measure combining data from three axes of motion) for a maximum of 9 hours at 1-s epochs, which limits the use of shorter epochs for this model.

One study by Mahar et al. (2008) has explored the epoch effect with pre-school children, where Actigraph data collected in 1-s epochs from 72 children (mean (SD) age: 3.9 (0.6) y) over 9 hours were reintegrated into 3-, 5-, 15-, 30- and 60-s epochs. The authors reported that using longer epochs resulted in significantly fewer minutes of moderate, MVPA, and vigorous intensity physical activity (p < 0.01). However, details of the study, including the cut-points used, which model of Actigraph accelerometer was used, and the resulting time spent at each intensity, were not provided. This limits the ability to determine whether the time differences would be considered biologically meaningful.

A large scale study by Edwardson & Gorely (2010), where accelerometry data were collected over 7 days from 311 children (7 - 11 y), also reported a significant epoch effect, between the means of the total number of minutes spent in MVPA between all epochs (5-s reintegrated into 15-, 30- and 60-s epochs). Interestingly, a shorter epoch length resulted in less time in MVPA, moderate (MPA) and light intensity physical activity (LPA) but more time in vigorous intensity physical activity (VPA). These differences, however may be partly explained by the use of different cut-points between studies, which makes cross-comparison difficult. The methodological issue of different cut-points will be discussed in section 1.6.4.

McClain et al. (2008) highlighted the need to consider the interaction between cut-points and epoch lengths. Estimates of physical activity were compared against the direct observation tool Computerized System for Observing Fitness Instruction Time (C-SOFIT) in 32 children (mean (SD) age: 10.3 (0.5) y) during a 30-minute Physical Education class. Cut-points from Treuth et al. (2004a), Mattocks et al. (2007) and Freedson et al. (2005) were applied to data collected in 5-s epochs reintegrated into 10-, 15-, 20-, 30- and 60-s epochs. While significant differences were reported between all epochs and C-SOFIT for the Treuth et al. (2004a) and Mattocks et al. (2007) cut-points, no significant difference was found for the Freedson et al. (2005) cut-points and the C-SOFIT for all epochs. While the outcome between epochs using the Freedson et al. (2005) cut-points were reported as being comparable, shorter epochs resulted in less individual error in estimates of MVPA.

In a more recent study of 86 children aged 4 - 10 y, Ojiambo et al. (2011) also reported that choice of epoch and cut-point significantly affected the classification of sedentary time and MVPA in data collected over a 6-day period. Importantly, when applying the Pate et al. (2006) cut-points, epoch length influenced the percentages of children seen to be complying with guidelines for MVPA, from 63% to 84% with 60-s and 15-s epochs respectively.

Finally, although some studies have divided the cut-point values developed from calibration studies using longer epochs to apply these to data collected in shorter epochs (Nilsson et al. 2002), McClain et al. (2008) cautions against the use of epoch adjusted cut-points until their validity has been established (McClain et al. 2008). However, a recent study by Jimmy et al. (2012) compared the effect of epoch on cut-points with data collected from 22 children aged 4 - 9 y (mean (SD) age: 6.71 (1.4) y). The authors concluded that the epoch does not seem to influence the cut-point values and that cut-points developed from longer epoch studies could be used in studies that apply a shorter epoch, through division by the appropriate factor. Despite this, there is limited evidence of the influence of epoch in studies of pre-school children and the implications of different epochs for the validity of cut-points is unknown.

In summary, studies of school-aged children suggest that shorter epochs improve sensitivity for detecting of VPA and very vigorous physical activity (Dorsey et al. 2009). While differences have been reported with the use of different epochs the biological significance of these has yet to be determined (Cliff et al. 2009b). In addition, whether any differences are apparent with pre-school children has not been established and which epoch or cut-points are most appropriate have yet to be determined (Ojiambo et al. 2011). Research is therefore needed to determine whether shorter epochs offer a more accurate means of quantifying physical activity levels in pre-school children (Welk et al. 2000b).

The first aim of this thesis was to determine whether shorter epochs offer a more accurate means of quantifying physical activity levels in pre-school children. This will bring new information to the field on the optimum epoch to use in studies which seek to explore habitual physical activity in pre-school children. A summary of the research question, aims and objectives are given in Table 1.9

Table 1.9: Research question and aims: studies of epoch effect.

Questions:

- What are the implications of shorter epochs on estimates of physical activity and sedentary behaviour in pre-school children (i.e. is there an ‘epoch effect’)?
- Which epoch is most accurate for measurement of physical activity in pre-school children?

Aims:

- To examine the influence of epochs on estimates of MVPA and sedentary behaviour of pre-school children with 7 - 10 days of free-living accelerometry data (Chapter 3).
- To determine which epoch is most accurate for measuring physical activity in pre-school children during free-play (Chapter 4).

1.6.3 Are there advantages to using triaxial accelerometry?

As discussed earlier in this chapter, different models of accelerometer exist which can measure acceleration of the body in one (vertical), two (vertical and mediolateral) or three (vertical, anteroposterior, mediolateral) planes (Vanhees et al. 2005). In addition, omnidirectional accelerometers are available which measure acceleration in any direction but are most sensitive in the vertical plane (Chen and Bassett 2005). The majority of studies with pre-school children to date have used uniaxial accelerometers (Taylor et al. 2009; Vale et al. 2010). The accuracy of studies which have used uniaxial accelerometers have been called into question as they are unable to detect movement in three dimensional planes (Trost et al. 2005) and debate exists about whether triaxial accelerometers would provide a more accurate estimate of pre-school children's physical activity (Cliff et al. 2009b).

There is some evidence to suggest that triaxial accelerometers may provide a more accurate estimate of physical activity in children compared to uniaxial models (Louie et al. 1999; Ott et al. 2000; Ott et al. 2000; Welk 2005). However, the differences in output between models are usually small and the correlation in the output is high which could indicate that similar information is being gathered (Rowlands and Eston 2007; Trost et al. 2005). Therefore despite the theoretical advantages of triaxial accelerometers over uniaxial monitors the benefits have not been established (Sirard and Pate 2001) and in addition this has not been adequately investigated with pre-school children (Cliff et al. 2009b).

One study by Kelly et al. (2004) compared both the uniaxial 7164 accelerometer and the omnidirectional Actiwatch (Mini Mitter Co., Inc., Bend OR) accelerometer against the direct observation tool the Children's Physical Activity Form (CPAF) in 78, 3- to 4-year-old children. The results suggested that the uniaxial accelerometry was significantly correlated against direct observation (r = 0.72, p < 0.001), while the Actiwatch was not (r = 0.16 p > 0.05). While these findings suggest no advantage of an omnidirectional accelerometer, the results should be viewed with caution given the limited validity of the CPAF. As already discussed, there are also limitations with using correlation methods for testing validity, which can fail to detect systematic differences between variables (Bland and Altman 1986).

To date there have been no studies which have explored whether there are advantages to using triaxial accelerometers to measure physical activity behaviour in pre-school children. The second aim of this thesis was to determine whether there were advantages to using triaxial accelerometry to measure physical activity behaviour of pre-school children. A summary of the research question, aims and objectives are given in Table 1.10.

Table 1.10: Research question and aim: study to compare triaxial versus uniaxial accelerometry.

Question:
• Are there advantages to using triaxial over uniaxial accelerometry to measure physical activity in pre-school children?
Aim:
• To investigate whether there are advantages to using triaxial over uniaxial accelerometry to measure physical activity in pre-school children during free-play (Chapter 5).

1.6.4 Which cut-points are most accurate?

Methodological issues remain around how accelerometry count data should be processed and interpreted to provide meaningful information needed for physical activity research (Cliff et al. 2009b).

Usually this is achieved through a process of calibration, whereby accelerometry counts are calibrated against a criterion measure, in order to convert the count output into a biologically meaningful format (Freedson et al. 2005). Methods such as indirect calorimetry or direct observation have been used as criterion measures, against which a range of count thresholds for different intensities of physical activity have been established (Freedson et al. 2005). However, the problem is that a number of calibration studies have been undertaken and a wide and divergent range of cut-points have been proposed and there is no consensus as to which cut-point to use (Ekelund et al. 2011; Reilly et al. 2008).

One problem is that many studies have calibrated cut-points against energy expenditure using linear regression analysis to establish ranges of accelerometer counts (cut-points) (Freedson et al. 2005). It is argued that there are issues with calibrating accelerometry data, which is a biomechanical measure of physical activity with energy expenditure and oxygen consumption, which are biological measures of physical activity (Freedson et al. 2005). In particular there are challenges with using biological measures of physical activity with children where the metabolic cost of movement, expressed relative to body mass ($ml \cdot kg^{-1} \cdot min^{-1}$), decreases as children mature (Davis 1980). Accelerometers have frequently been validated against indirect calorimetry and calibrated in terms of resting metabolic equivalents (METs). However, as discussed in section 1.3.1 (p 2), the use of 1 MET as being 3.5 $ml \cdot kg^{-1} \cdot min^{-1}$, as used in adult studies, is not applicable for use with children as their resting metabolic rates (RMR) are much higher. Using adult measures to represent RMR can therefore introduce systematic error and the cut-points derived from adult studies are not appropriate to apply to data collected from children (Treuth et al. 2004a). To ensure accuracy it is important that calibration studies include a measure of a child's RMR so that these can be used to estimate the child's own MET values. While this individual calibration approach is useful it possibly limits the feasibility of using accelerometers to estimate EE in large population-based studies.

According to Freedson et al. (2005), calibration of activity counts using behavioural approaches, such as direct observation methods, offers an alternative to biological calibration methods which are useful when researchers are interested in patterns and determinants of physical activity. It is argued that observation systems avoid interpretation errors associated with METs and errors associated with extrapolation from treadmill activity to free-living behaviours. They may also be particularly useful with studies of young children where measurement and interpretation of energy expenditure can be difficult (Freedson et al. 2005; Oliver et al. 2007b).

1.6.4.1.1 Calibration studies of Actigraph cut-point for children

Table 1.11 presents a summary of calibration studies for the Actigraph accelerometer in school-aged children and adolescents (children aged > 5 to 18 y). To allow cross comparison between studies the cut-points published as 1-min epochs have been divided to provide a value for 15 s (Nilsson et al. 2002). The reason for considering calibration studies conducted with school-aged children is that many of the cut-points developed from these studies have been applied in studies of pre-school children and there is a debate over whether age specific cut-points are required (Mackintosh et al. 2012) or not (Evenson et al. 2008). A range of different cut-points have been proposed and these may be as a result of differences in criterion measures as well as differences in the protocol used in calibration studies. Some studies have made use of protocols which include treadmill-based activities (Dowda et al. 1997; Freedson et al. 1997) and others include structured or unstructured activities (Mattocks et al. 2007). In addition the criteria used for determining different intensities of activity and the activities included within the different categories of intensity of activity, vary between studies. For example, some studies include standing as a sedentary activity (Sirard et al. 2005), while others categorise standing as a light intensity physical activity (Trost et al. 2012).

Puyau et al. (2004), calibrated cut-points for the 7164 accelerometer using a room calorimeter (n = 26, age: 6 to 16 y). Children and adolescent participants were asked to undertake structured activities that included treadmill walking and running. While MET levels were calculated in terms of the child’s individual RMR, they were not used to define thresholds for intensity. These were instead defined in terms of activity energy expenditure (AEE), calculated as EE - RMR. Light activity was defined as activities involving low level exertion in the standing position, moderate activities involved medium exertion in the standing position and vigorous activity involved high level exertion in the standing position. It was found that while the relationship between EE and counts was dependent on age, the relationship between AEE and counts was independent of age and sex, with age not significantly altering the prediction of AEE from counts. A linear regression of AEE on activity counts was used to define threshold counts for sedentary, LPA, moderate and vigorous activity. There are, however, known issues with the use of linear regression equations, as aerobic energy expenditure is often non-linear in nature at its upper levels due to the contribution of anaerobic energy sources (Freedson et al. 2005). In addition, during intermittent free-play type activities the relationship between counts and energy expenditure is also not thought to be linear (Matthews 2005). Finally, there are problems with linear regression equations used in the calibration of Actigraph accelerometers, where counts at higher intensity have been shown not to increase linearly (Brage et al. 2003).

Freedson et al. (1997) calibrated cut-points for children (n = 80, age: 6 to 17 y) during walking and running on a treadmill. The calculation was made using accelerometer counts and respiratory gas exchange through a minute-by-minute analysis. A limitation of this study was that it was based on treadmill activity only and did not include any of the free-living activities which are typical of children.

In another calibration study, Dowda et al. (1997) calibrated the 7164 accelerometer cut-points (n = 80, 6 to 18 y) in children against indirect calorimetry during treadmill walking and running. In the study the researchers calculated count thresholds for a given MET value. The authors used pre-defined MET thresholds for moderate activity as three to 5.99 METs, vigorous as six to 8.99 METs and very vigorous physical activity as more than or equal to nine METs. When exploring different treadmill walking speeds the authors identified age-specific count thresholds which corresponded to the pre-defined METs values (see Table 1.11). It was interesting to note that even at the slowest treadmill speeds children below the age of 12 had an energy expenditure of ≥ 3.8 METs. A limitation of this study is that details of how the MET values were calculated, and if these were based on individual RMR, are unclear.

There is concern over the use of three METs as a threshold for moderate intensity physical activity in children. This was highlighted in a study by Treuth et al. (2004a), where the 7164 accelerometer was calibrated against heart-rate and indirect calorimetry in 74 females aged 13 to 14 years (mean (SD) age: 14.1 (0.3) y). Based on each child's RMR the authors identified the lower end of the moderate intensity range as approximately 4.6 METs. A limitation of this study is that the authors used two walking speeds, slow and brisk walk, to distinguish moderate from LPA and this was therefore not based on 'free-living' activities.

However, Mattocks et al. (2007) reported similar findings in their study, in which the Actigraph accelerometer was calibrated against indirect calorimetry in 246 children (mean (SD) age: 12.4 (0.2) y) while undertaking a series of structured, 'free-living' indoor activities. The authors argued that 4 METs may be a more appropriate threshold for moderate intensity physical activity in children (Mattocks et al. 2007).

An alternative to regression analysis is to use receiver operating characteristic curve (ROC) analysis where cut-points can be defined based on desired sensitivity and specificity (Trost et al. 2011). ROC analysis involves plotting a graph of the relationship between the true positive rate (sensitivity) and the false positive rate (specificity) analysis (Joy et al. 2005). One of the few calibration studies to use ROC curve analysis in the calibration of cut-points is by Evenson et al. (2008), who determined thresholds for the Actigraph accelerometer for 33 children aged 5 to 8 y (mean (SD) age: 7.3 (1.1) y). Using ROC analysis, cut-points identified from plotting sensitivity and specificity curves were used to determine thresholds for sedentary, moderate and vigorous activities. This study determined each individual child's resting $VO_{2,}$ which as earlier studies have stated is important for accuracy in studies of children (Freedson et al. 2005). The researchers compared the threshold values for 5 to 6 year olds versus 7 to 8 year olds for each intensity level and there was no significant difference except for vigorous activity and the authors suggest that age-specific cut-points are not required. Interestingly this study included static cycling and it could be seen that despite having the 6th highest VO_2 value (19.3 ml $kg^{-1}{\cdot}min^{-1}$), the accelerometry counts for this were similar to sedentary activity (164 cpm), which highlights the limitation of accelerometers in accurately measuring certain activities.

A limitation of many of the calibration studies is that they are laboratory based studies and have calibrated cut-points during either treadmill or structured activities, which may not reflect young children's activity, which is not usually performed in sustained bouts (Riddoch et al. 2007). Applying cut-points which have been calibrated against structured activities to free-living activities may result in intermittent physical activity being misclassified as sedentary behaviour (Mackintosh et al. 2012; Welk 2005). It has been argued that in order to develop behaviourally valid cut-points, 'field-based' protocols, which incorporate free-living activities without controlling a child's behaviour, should be adopted during calibration studies (Mackintosh et al. 2012). In addition, according to Ott et al. (2000), measurement tools should be calibrated for use with activities which involve a range of activities, such as bending, jumping and running, which are characteristic of the activities of children.

Mackintosh et al. (2012) calibrated a population-specific protocol for use with school-aged children (10 - 11y) using the SOPLAY direct observation scale as a criterion measure. This study provides a useful protocol for calibrating accelerometers which incorporates free-living indoor and outdoor activities and is suitable for use in school-aged children. However, this protocol may not be suitable for younger child as developmentally they may not be able to engage in games such as Frisbee, and Hopscotch for a sustained period of time.

Table 1.11: Calibration studies of Actigraph accelerometry cut-points for children > 5 years of age.

Authors	No. participants (n), age range, mean (SD)	Criterion	Activities	Counts/15 s Sed	LPA	MVPA
Dowda et al.(1997)	80, 6 - 18 y	VO_2 (portable metabolic unit) MVPA ≥ 3 METs	Treadmill W,R	N/A	N/A	12 y ≥ 131 18 y ≥ 537
Eston et al. (1998)	30, 8.2 - 10.8 y	VO_2 (portable metabolic unit)	W,R,FL	N/A	N/A	> 125
Evenson et al. (2008)	33, 5 - 8 (1.1) y	VO_2 (portable metabolic unit)	Structured activities Treadmill W,R	≤ 25	> 25 to ≤ 573	≥ 574
Freedson et al. (2005) (based on 12 y old Trost et al. (2011))	80, 6 - 18 y	VO_2 (portable metabolic unit) MVPA ≥ 3 METs	Treadmill W, R	≤ 25	> 25 to < 555	≥ 555
MacKintosh et al. (2012)	28, 10 - 11 y	SOPLAY ROC analysis	Drawing; watching DVD; W,R, Games	≤ 93	> 93 to ≤ 540	> 540
Mattocks et al. (2007)	246, 12.4 (0.2) y	VO_2 (portable metabolic unit) MVPA ≥ 4 METs	Structure activities, W,R	N/A	N/A	≥ 895
Metcalf et al. (2008)	NA unpublished	NA MVPA ≥3 METs =~4km·h^{-1}	Treadmill W	N/A	N/A	≥ 625
Puyau et al. (2002)	26, 6 - 16 y	VO_2 (whole room calorimetry)	FL	< 200	≥ 200 to < 800	≥ 800
Pulsford et al. (2011)	53, 7 - 8 y	VO_2 (portable metabolic unit)	Structured activities, W, R	< 25	≥ 26 to ≤ 560	> 561
Treuth et al. (2004a)	74, 13 - 14 y; Girls only	VO_2 (portable metabolic unit) MVPA ≥ 4.6 METs	Structured activities, Treadmill W,R	≤ 25	> 25 to <750	≥ 750
Vanhelst et al. (2010b)	40, 10 - 16 y	VO_2 (portable metabolic unit)	Structured FL, and treadmill	< 100	≥ 101 to < 476	≥ 476

FL: free-living; LPA: light physical activity; METs: Metabolic equivalent units; MVPA: moderate-to-vigorous physical activity; R: running; Sed: Sedentary behaviour; SOPLAY: System for Observing Fitness Instruction Time; W: walking.

1.6.4.1.2 Calibration studies of Actigraph cut-points for pre-school children

There are a growing number of studies which have calibrated cut-points in younger children and these are summarised in Table 1.12. Similar to school-aged children, there is a range of proposed cut-points for the different intensities of physical activity and sedentary behaviour.

Pate et al. (2006) was the first study to provide calibration and cross-validation data using indirect calorimetry as a criterion measure in young children. In their study of 29 children (3 to 5 y), they identified the threshold of 20 ml $kg^{-1}\cdot min^{-1}$ to differentiate between slow and brisk walk, which they classified as moderate intensity physical activity, corresponding to activity counts of > 420 counts/15 s. This threshold of counts for moderate activity is lower than those developed from an earlier calibration study by Sirard et al. (2005). This might in part be explained by a different criterion measure being used in the calibration process. Sirard et al. (2005) made use of the direct observation tool the CARS (Puyau et al. 2002) as their criterion method.

Reilly et al. (2003) was the first to publish cut-points for sedentary behaviour for pre-school children. In their study of thirty, 3 to 4 year olds (mean (SD) age: 3.7 (0.5) y) children were observed for an average (SD) of 100 (17) min while wearing the 7164 accelerometer, which they compared against the CPAF (O'Hara et al. 1989). Using ROC analysis, the optimal cut-point for accelerometry output for sedentary behaviour was identified. The identified cut-point of < 1100 cpm for sedentary activity in the cross-validation had a mean specificity of 83% and a mean sensitivity of 82%. Recent research with older children is advocating the use of a much lower cut-point for sedentary behaviour (< 100 cpm) (Evenson et al. 2008) and suggests that using 1100 cpm may result in some light intensity physical activity being incorrectly classified as sedentary behaviour.

Table 1.12: Calibration of studies of Actigraph accelerometry cut-points for pre-school children (3 to 5 years).

Authors	No. participants (n), age range, mean (SD)	Criterion	Activities	Counts/15 s		
				Sed	LPA	MVPA
Jimmy et al. (2012)	22, 4 - 9 y	VO_2 (portable metabolic unit)	Rest, W,R, FP	NA	NA	≥ 429
Pate et al. (2006)	29, 3 - 5 y	VO_2 (portable metabolic unit)	Rest, W,R, FP	NA	NA	≥ 420
Reilly et al. (2003)	30, 3 - 5 y	CPAF	FP	< 275	NA	NA
Sirard et al. (2005)	16, 3- , 4-, 5- year olds	CARS	Structured activities, W, R, FP			
			3 y olds	< 301	≥ 302 to ≤ 614	≥ 615
			4 y olds	< 363	≥ 364 to ≤ 811	≥ 812
			5 y olds	< 398	≥ 399 to ≤ 890	≥ 891
Van Cauwenberghe et al. (2011)	18, 5.8 (0.3) y	CARS	FP, Structured activities, treadmill W, R	< 373	≥ 373 to < 585	≥ 585

CARS: Children's Activity Rating Scale; CPAF: Children's Physical Activity Form; FP: free-play; LPA: light physical activity; MVPA: moderate-to-vigorous activity; NA: not available; R: running; Sed: sedentary behaviour; W: walking.

1.6.4.1.3 Studies comparing cut-points in pre-school children

Studies have highlighted the problem of applying different cut-points to accelerometry data in estimates of time spent in different intensities, in particular for time spent in MVPA and TPA (Beets et al. 2011b; Cliff and Okely 2007; Guinhouya et al. 2006; Kim et al. 2012; Mota et al. 2007). The lower cut-points for MVPA proposed by Freedson et (1997), Pate et al. (2006) and Trost et al. (2002) have been compared against the higher cut-points for MVPA published by Sirard et al. (2005) and Puyau et al. (2002). Prior to 2011, this had substantial implications when considering whether children were meeting the then recommendations for 60 minutes of MVPA per day and conflicting conclusions were reached.

Guinhouya et al. (2006) compared the cut-points defined by Trost et al. (2002) with those defined by Puyau et al. (2002) in accelerometry data collected from 45 children aged 8 to 11 years over a 3-day period. The authors reported a large significant difference in time spent in MVPA with mean (SD) minutes of MVPA of 141 (39) min/day for the Trost et al. (2002) cut-points and 28 (18) min/day for the Puyau et al. (2002) cut-points. Guinhouya et al. (2006) concluded that the Puyau et al. (2002) cut-points were more discriminatory, while the Trost et al. (2002) cut-points reduced the gap between the most and least active child, which may result in a greater risk of type I errors occurring as a consequence.

In another study, Mota et al. (2007) compared the Freedson et al. (1997) and the Puyau et al. (2002) cut-points using data collected from 63 children (8 to 15 y) over 3 days. Discrepancies in time spent in MVPA were reported, with 95% of boys and 87.2% of girls seen to be meeting the recommendation of 60-minutes of MVPA per day with the Freedson et al. (1997) cut-points and only 17.4 % of boys and 5.1% of girls using the Puyau et al. (2002) cut-points. The authors argued that the high number of children seen to be meeting the recommendation of 60 min MVPA per day when the Freedson et al. (1997) cut-points were applied was in agreement with other studies (Hussey et al. 2001; Sleap and Tolfrey 2001). However it should be noted that one of the supporting studies by Hussey et al. (2001) is based on data collected from a proxy questionnaire (Aaron et al. 1995). Not only are questionnaires a subjective measure of activity, but as discussed, there are known issues with overestimation in reported physical activity using proxy reports (Bender et al. 2005).

Even cut-points which are quite similar can be problematic when applied to data. This was demonstrated in a study by Cliff and Okely (2007), who processed accelerometry data collected from 58 children aged 3 to 5 years (mean (SD) age: 4.38 (0.8) y) over 5 days, using the Puyau et al. (2002) and the Sirard et al. (2005) cut-points. While there was no significant overall difference between mean number of minutes of MVPA, there were, however, large limits of agreement (LOA) (-28.04 to 28.94 min/day). Also, when the sample was categorised into age bands, there was a large difference between the mean number of minutes of MVPA, depending on which cut-point was applied. This was most apparent for the 3 year olds, where a mean error of 32.35 min/day was reported. The difference for 3 year olds of more than 30 min/day could have been exacerbated by the small sample size of 3 year olds (n = 8) included in this study.

There is only one study, by Reilly et al. (2008), which has compared several cut-points for time spent in MVPA and sedentary activity, in pre-school children (n = 72; mean (SD) age: 5.8 (0.5) y). The authors applied the commonly used cut-points by Puyau et al. (2002), Freedson et al. (1997), Treuth et al. (2004a) and Reilly et al. (2003) to data collected from children over a 7 day period. There were large differences with the Freedson et al. (1997) cut-points, resulting in 266 min (nearly 4 hours) of MVPA per day and the Puyau et al. (2002) and Treuth et al. (2004a) cut-points resulting in an estimated 28 min and 41 min of MVPA per day, respectively. The authors question whether it is plausible that young children are engaging in 4 hours of MVPA per day given the increasing trends of obesity and the evidence from studies using heart-rate, direct observation and pedometry, which suggests that levels of MVPA are low (Cardon and De Bourdeaudhuij 2007; McKee et al. 2005).

Several factors in the calibration process may explain why different cut-points have been proposed. Studies have used different criterion methods, different epoch lengths and, as already stated, it is unclear whether cut-points calibrated at 1-min epochs are valid at 15-s epochs. In addition, the protocols for activities used in calibration are not always reflective of spontaneous free-living activities (Baquet et al. 2007). Finally, there are concerns regarding whether different generations of Actigraph are comparable (Ried-Larsen et al. 2012) and it is unknown whether it is appropriate to use cut-points developed for the 7164 accelerometer for the later generations of Actigraph accelerometers, for example, the GT1M and GT3X models.

The fact that currently a range of cut-points exist highlights a lack of agreement regarding interpretation of data (Rowlands 2007). Cliff and Okely (2007) argue that more rigorous testing of pre-school children's cut-points are necessary to reach greater consensus on the most appropriate cut-points to use. While studies have compared the consequences of applying different cut-points, to date, no study has compared the results against a criterion measure with pre-school children to see which cut-points are most accurate. It is also important to establish if these cut-points are valid for free-living activities in young children. The aim of this study was therefore to determine which cut-points were most accurate to categorise physical activity intensity and sedentary behaviour during free-play with pre-school children. Table 1.13 summaries the aims of the thesis.

Table 1.13: Research question and aim: study of cut-points.

Question: • Which Actigraph accelerometry cut-points are most accurate for pre-school children? **Aim:** • To validate Actigraph accelerometry cut-points for estimating physical activity and sedentary behaviour in pre-school children during free-play (Chapter 6).

1.6.5 Are different generations of Actigraph accelerometers comparable?

Actigraph accelerometers are among the most widely utilised motion sensor in children's physical activity research (Trost et al. 2011). The latest generation of Actigraph accelerometers (GT1M and GT3X) have a different internal technology, using a Micro-Electro-Mechanical System (MEMS) solid state accelerometer as opposed to the piezoelectric cantilevered beam accelerometer, seen in the older generation model 7164 (John et al. 2010). In addition, while both generations of monitor capture time varying acceleration in the range of 0.05 to 2 G, there are differences in the filtering and sampling frequencies between models. As a result of these differences, the same acceleration could result in different count output (Chen et al. 2012) and the two generations of monitors cannot necessarily be used interchangeably.

Comparison studies between the GT1M and the 7164 have been undertaken using a mechanical set-up (Rothney et al. 2008), during treadmill activities with adults in laboratory settings (Fudge et al. 2007; John et al. 2010; Kozey et al. 2010b), and in a free-living situation with adolescents (Corder et al. 2007).

Rothney et al. (2008) subjected GT1M and 7164 accelerometers to a range of accelerations by simultaneously oscillating the accelerometers using a mechanical set up. The authors reported significant differences in activity counts between models, however, concluded that these differences would likely have a minimal impact on time spent at different intensities (Rothney et al. 2008).

The findings from laboratory based studies which have compared the different generations of accelerometers with participants are inconclusive. John et al. (2010) compared the accelerometry models in a study of 10 male participants (mean (SD): 23.6 (2.7) y), who completed treadmill walking and running at 10 different speeds while wearing the 7164 and the GT1M accelerometers. The researchers reported as a non significant difference of (mean (SD)) 439 (982) counts per minute (cpm) between the different generations of Actigraph accelerometer. In contrast, in a study of 16 subjects (mean (SD) age: 23 (3) y) which used a similar study protocol, Fudge et al. (2007) reported higher output from the GT1M compared with the 7164 accelerometer. Similarly, Kozey et al. (2010b) reported 2.7% higher count output of the GT1M compared to the 7164 accelerometer across three self-selected speeds on a treadmill: slow 0.7 (0.2) $m{\cdot}s^{-1}$, medium 1.3 (0.2) $m{\cdot}s^{-1}$ and fast 2.1 (0.6) $m{\cdot}s^{-1}$ (mean (SD)).

Despite significant differences between model output ($p < 0.05$) the authors concluded that these differences would not result in any meaningful differences in intensity classifications.

Only one study to date has compared the GT1M and 7164 during free-living conditions with 30 Indian adolescents (mean (SD) age: 15.8 (0.6) y) over a 7 day continuous monitoring period. The findings suggested that the GT1M counts were on average 9% lower ($p < 0.05$) than those derived from the 7164 accelerometer. The authors recommended either applying a correction factor of 0.91 to the GT1M data, or reducing the 7164 cut-points by 10% for the GT1M data to allow comparison between models (Corder et al. 2007). In this study the differences did not translate into observable differences in time spent in moderate or vigorous physical activity. Whether applying this correction factor, or altering the cut-points for GT1M data would allow comparison between accelerometry models has not been tested and has not been explored in studies of children.

Given that several large scale longitudinal studies are reporting on the tracking of physical activity (Kristensen et al. 2008; Metcalf et al. 2008; Mitchell et al. 2012), it is important to ensure that systematic measurement errors have not taken place as a consequence of using different generations of accelerometers. There are also implications of applying cut-points calibrated from an older generation to a newer generation of accelerometer, and it is unclear whether the cut-points are valid for different generations of accelerometers. Investigation of the comparability between output from different generations of accelerometers used in longitudinal studies is therefore essential to ensure that population-based estimates of physical activity are unbiased (Ried-Larsen et al. 2012). In a recent article, Ried-Larsen et al. (2012) stressed the need for more free-living studies in different populations in order to interpret the observed differences seen in mechanical calibration studies and to explore whether any differences between monitors are population specific.

Comparability between different generations of accelerometers has not previously been explored in children or in young children. This is the fourth aim of this thesis, which is presented in Table 1.14.

Table 1.14: Research question and aims: study comparing different generations of Actigraph.

Question:

- Are different generations of Actigraph accelerometer comparable when used with pre-school children?

Aims:

- To compare different generations of Actigraph accelerometer during mechanical calibration (Chapter 7).
- To compare different generations of Actigraph accelerometer in pre-school children during 1 hour of free-play (Chapter 7).

1.6.6 What is the recommended wear time to provide a reliable estimate of physical activity and sedentary behaviour in pre-school children?

There are methodological questions around the number of days of data collection needed and the number of hours to constitute a 'valid' day, to accurately capture habitual physical activity (Hinkley et al. 2012b). There is some evidence that as few as 3 hours a day can provide reliable estimates of physical activity for the pre-school age group and with the difference between 3 and 10 hours being minimal, when 1 to 7 days of monitoring are used (Penpraze et al. 2006). While 10 days of monitoring maximises reliability, estimates suggest that 3 days may have sufficient reliability for many purposes ($r \geq 0.60$) (Penpraze et al. 2006). More recently, published data from children aged 4 to 10 y (mean (SD) age: 7 (2) y) suggest that a minimum duration of 6 hours for 7 to 9 days of monitoring including one weekend day, is necessary to achieve 80% reliability (Ojiambo et al. 2011). However, increasing the number of days of monitoring brings the risk of decreasing compliance and therefore a compromise has to be reached. While earlier studies suggest that there are no differences in weekend and week days for young children (Sigmund et al. 2007), more recent studies suggest that differences may exist in activity levels between weekday and weekend days (Hinkley et al. 2012c; Vale et al. 2010). It is not clear whether the inclusion of a weekend day is necessary in analysis to obtain a reliable estimate of habitual physical activity.

Another area which may influence estimates of physical activity and sedentary behaviour is how the accelerometry data are cleaned and reduced (Masse et al. 1999; Mâsse et al. 2005), for which a variety of methods have been adopted (Rowlands 2007). As part of the data cleaning process, researchers need to make decisions about excluding or imputing data where there are extended periods of consecutive zeros. This is often referred to as 'non-wear' time and it is important to distinguish between sedentary time when the monitor is worn but the person is stationary and 'true' non-wear time when the monitor is removed. These episodes may look similar in the data. Criteria have been proposed to exclude non-wear time from analysis and these range from excluding 10 min of consecutive zeros, to 180 min of consecutive zeros being excluded (Oliver et al. 2012; Rowlands 2007). Cliff et al. (2009b) have reviewed the literature and made recommendations for dealing with non-wear time and for screening data for upper limits of biological plausibility. However, the evidence for these recommendations is based on studies of older children (Esliger et al. 2005) and the influence that different criteria for non-wear time have on estimates of physical activity and sedentary behaviour has not been investigated in younger children

(Cliff et al. 2009b). The final aim of this thesis is to determine the recommended number of days and hours of data needed to provide a reliable estimate of habitual physical activity and sedentary behaviour in pre-school children. In addition, it is an aim to determine whether inclusion of a weekend day influences reliability. A final aim is to investigate the influence that applying different criteria for non-wear time has on estimates of habitual physical activity and sedentary behaviour. The aims are summarised in Table 1.15

Table 1.15: Research questions and aims: study of wear time.

Question:

- What is the recommended wear time to provide a reliable estimate of habitual physical activity and sedentary behaviour in pre-school children?

Sub questions:

  - How many days and hours of data are needed for stable measurement of physical activity and sedentary behaviour?
  - Are weekend days necessary for reliable estimates of habitual physical activity and sedentary behaviour?
  - Does the application of non-wear time criteria (missing data) influence estimates of physical activity and sedentary behaviour in pre-school children?

Aims:

- To determine the recommended wear time required to provide a reliable estimate of habitual physical activity and sedentary behaviour of pre-school children (Chapter 8).
- To investigate whether the inclusion of a weekend day is necessary for reliable estimates of habitual physical activity and sedentary behaviour of pre-school children (Chapter 8).
- To examine the influence of applying non-wear time criteria to estimates of physical activity and sedentary behaviour of pre-school children with 7 days of free-living accelerometry data (Chapter 8).

1.6.7 Summary

Accurate means of quantifying physical activity are important for gaining an insight into the physical activity behaviours in populations. Accelerometry offers a valid and reliable means of objectively measuring physical activity in pre-school children. Despite this, several major outstanding methodological questions remain regarding the use of accelerometers and the different approaches adopted in studies can impact on whether children are seen to be meeting recommendations for health or not. It is only by addressing the discrepancies in methodological decisions that accurate means of measuring physical activity can be strived for. This is fundamental for future research which seeks to quantify physical activity levels and to understand the relationship between physical activity and health and for evaluating interventions designed to promote physical activity (Trost 2007).

Table 1.16 presents a summary of the questions on which this thesis is based, and the chapters in which these will be addressed.

Table 1.16: Summary of questions on which this thesis is based.

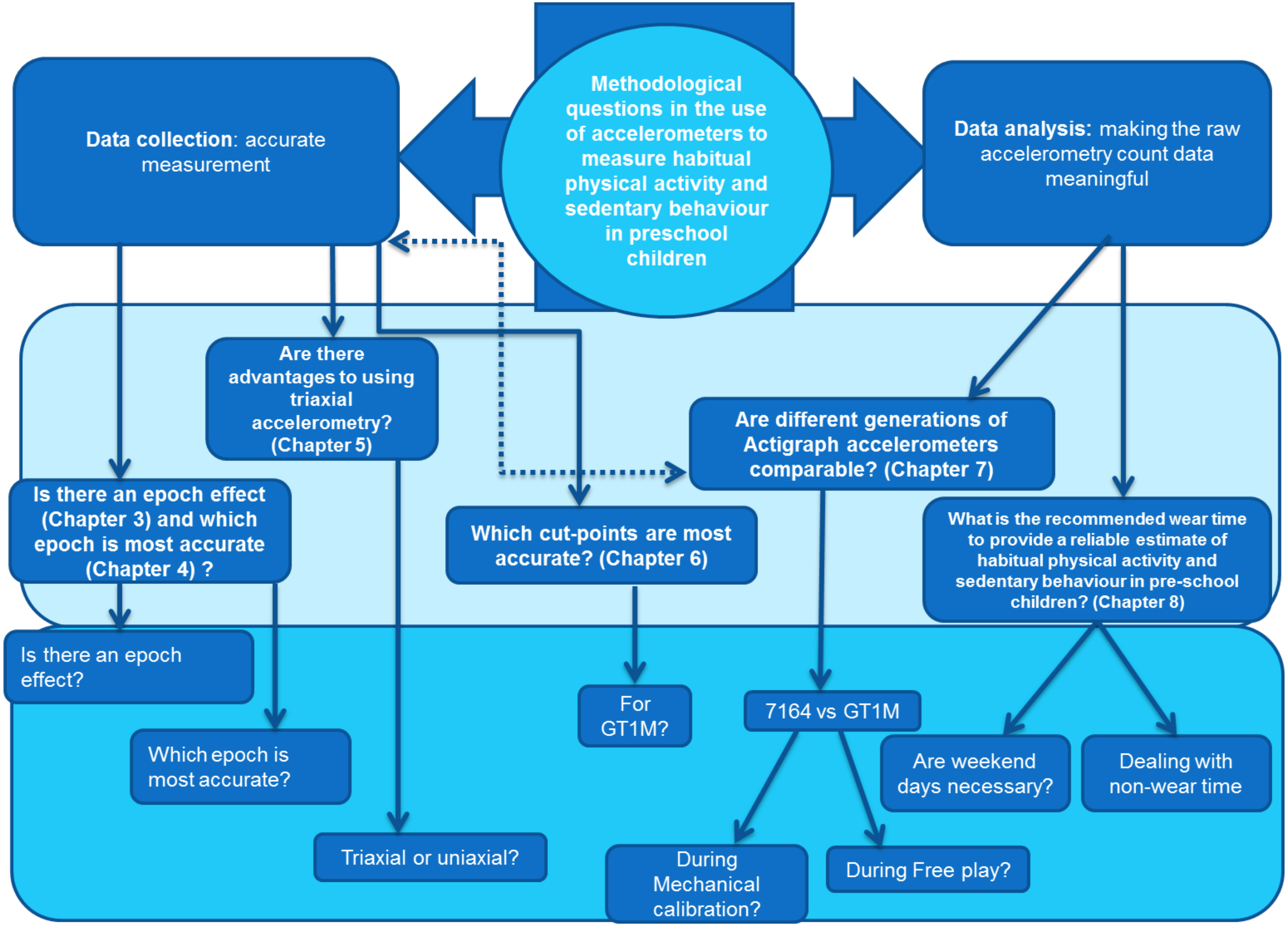
Research questions	Evidence	Chapter
What are the implications of shorter epochs on estimates of physical activity and sedentary behaviour in pre-school children (i.e. is there an 'epoch effect')?	Some evidence with school aged children that longer epochs result in high intensity physical activity being misclassified as lower intensity. (Nilsson et al. 2002; Rowlands et al. 2006)	See Chapter 3
Which epoch is most accurate for measurement of physical activity in pre-school children?	Lack of evidence in studies with pre-school children.	See Chapter 4
Are there advantages to using triaxial accelerometry to measure physical activity in pre-school children?	Theoretical advantage of triaxial accelerometry. No empirical evidence (de Vries et al. 2006; Oliver et al. 2007b; Oliver et al. 2009; Rowlands 2007; Ward et al. 2005).	See Chapter 5
Which Actigraph accelerometry cut-points are most accurate for pre-school children?	Wide variation in practice; implications of using different cut-points not appreciated.	See Chapter 6
Are different generations of Actigraph accelerometer comparable when used with pre-school children?	Possibly a difference seen in studies of adolescents (Corder et al. 2007). No empirical evidence from studies of pre-school children.	See Chapter 7
What is the recommended wear time to provide a reliable estimate of habitual physical activity and sedentary behaviour in pre-school children? • How many days and hours of data are needed for stable measurement of physical activity and sedentary behaviour? • Are weekend days necessary for reliable estimates of habitual physical activity and sedentary behaviour? • Does the application of non-wear time criteria (missing data) influence estimates of physical activity and sedentary behaviour in pre-school children?	• Recommendations range from 3 to 7 days (Cliff et al. 2009b; Hinkley et al. 2012b; Penpraze et al. 2006). • Recommendations range from 3 hours to 10 hours (Penpraze et al. 2006). • Unclear if weekend days are necessary or not (Hinkley et al. 2012b; Penpraze et al. 2006). • Lack of consistency in practice with criteria used for non-wear time, implications of this unclear.	See Chapter 8

Adapted from Reilly et al. (2008)

1.7 STRUCTURE OF THE THESIS

This thesis comprises the following chapters: Chapter 2 will outline the general methods adopted in chapters four to seven, which report on a 1-hour observational study. Chapter 3 presents the methods and findings of a study which investigated the influence that applying different epochs had on estimates of MVPA. This was a secondary data analysis of free-living accelerometry data collected from young children (n = 32, 17 males, 15 females; mean (SD) age: 5.9 (0.7) y) over 7 - 10 days. Chapter 4 presents findings from a 1-hour observational study of free-play of pre-school children (n = 31, 15 males, 16 females; mean (SD) age: 4.4 (0.8) y), in which the accuracy of different epochs are compared against direct observation. Chapter 5 provides a comparison of uniaxial versus triaxial accelerometry during 1 hour free-play of pre-school children (n = 31, 15 males, 16 females; mean (SD) age: 4.4 (0.8) y). Chapter 6 investigates the accuracy of accelerometry cut-points against direct observation in pre-school children (n = 31, 15 males, 16 females; mean (SD) age: 4.4 (0.8) y) during 1 hour of free-play. Chapter 7 compares the 7164 and GT1M Actigraph accelerometers during mechanical calibration and against direct observation during 1 hour of free-play with pre-school children (n = 23, 10 boys, 13 girls; mean (SD) age: 4.3 (0.8) y). Chapter 8 presents findings of an investigation into the recommended wear time to provide a reliable estimate of habitual physical activity and sedentary behaviour in free-living accelerometry data collected from pre-school children (n = 112, 60 males, 52 females, mean (SD) age: 3.7 (0.7) y) over 7 days. This study also investigates whether weekend days are necessary for reliable estimates of habitual physical activity in pre-school children and explores the influence of applying different criteria for non-wear time (e.g. excluding 10, 20 and 60 s of consecutive zeros). Chapter 9 presents a general discussion of the findings of the thesis and finally, Chapter 10 provides a conclusion to the thesis. An overview of the thesis is presented in Figure 1.2.

Figure 1.2: Overview of thesis.



1.8 SUMMARY OF AIMS

The aims of this thesis are:

- To examine the influence of epochs on estimates of MVPA and sedentary behaviour of pre-school children with 7 - 10 days of free-living accelerometry data (Chapter 3).

- To determine which epoch is most accurate for measuring physical activity in pre-school children during free-play (Chapter 4).

- To investigate whether there are advantages to using triaxial over uniaxial accelerometry to measure physical activity in pre-school children during free-play (Chapter 5).

- To validate Actigraph accelerometry cut-points for estimating physical activity and sedentary behaviour in pre-school children during free-play (Chapter 6).

- To compare different generations of Actigraph accelerometer during mechanical calibration (Chapter 7).

- To compare different generations of Actigraph accelerometer in pre-school children during 1 hour of free-play (Chapter 7).

- To determine the recommended wear time required to provide a reliable estimate of habitual physical activity and sedentary behaviour of pre-school children (Chapter 8).

- To investigate whether the inclusion of a weekend day is necessary for reliable estimates of habitual physical activity and sedentary behaviour of pre-school children (Chapter 8).

- To examine the influence of applying non-wear time criteria to estimates of physical activity and sedentary behaviour of pre-school children with 7 days of free-living accelerometry data (Chapter 8).

CHAPTER 2 : GENERAL METHODS

This chapter provides detail on the general methods and instrumentation used in the studies within this thesis. Specific details of individual studies are outlined in the relevant study chapters.

2.1 STUDY DESIGN

The studies conducted in chapters four to seven were based on a 1 hour cross-sectional observational study of pre-school children (aged 3 to 5 y). Children were video recorded during their usual time-tabled outdoor playtime in their nursery while they wore accelerometers. This allowed the researcher to observe and code the children's physical activity behaviour and compare this with simultaneously collected accelerometry data. Chapter 3 and Chapter 8 were cohort studies and involved secondary data analysis of free-living physical activity data collected from pre-school children (aged 3 to 5 y) over a 7-to 10-day-period. This current chapter will outline the ethical procedures, recruitment and the protocol adopted in the cross-sectional observational study. The ethical procedures, recruitment and protocol for the cohort studies are discussed in the respective chapters (Chapter 3 and Chapter 8). Finally, the accelerometry equipment used in all studies is discussed in this general methods chapter.

2.2 ETHICAL APPROVAL

Ethical approval for the observational study was granted from Queen Margaret University Ethics committee (Appendix II a). Permission to undertake the study was granted from Edinburgh City Council, Children and Families Department (Appendix II b). Two nurseries in Edinburgh, CGU5 and KLN were identified by the Children and Family Department and permission was obtained from the head teacher of each nursery to undertake the study on their premises. Information sheets and written consent forms were developed for parents and carers (Appendix II c and Appendix II d) and an age appropriate information pamphlet (Appendix II e) was developed for the parents to read to their child together with a verbal assent form (Appendix II f).

2.3 PARTICIPANTS AND RECRUITMENT

A convenience sample was selected by inviting pre-school aged children (aged 3 to 5 years) attending two Edinburgh city council pre-schools to take part. Inclusion criteria were that the child was apparently healthy, between the ages of 3 to 5 years, and with no known chronic disease relating to energy expenditure or physical activity (Fisher et al. 2005b). In addition, only those children who had given verbal assent and whose parents had given written consent for them to take part in the study were included. Children with any known physical problems that could affect their mobility, including neurological, respiratory or musculoskeletal problems, were excluded from the study.

To recruit participants, the nurseries distributed information sheets and consent forms to all parents of pre-school children attending their nursery, together with a stamped addressed envelope for return of a consent form to the researcher. A verbal assent form for the children was also included and this was for the parents to read to their child to obtain the child's permission to be involved in the study. The researcher also obtained verbal assent from the children at the time of data collection and the child could opt out from wearing the accelerometer at any point.

2.4 SAMPLE SIZE CALCULATION AND SAMPLE CHARACTERISTICS

Based on *a priori* power of 0.80, and an α level 0.05, the sample size calculation estimated 26 participants as being necessary for an effect size of 0.50, to detect a statistically significant correlation between the count output of two accelerometry models (Puyau et al. 2002). The paired design of the observational study (subjects acted as their own match) is argued to increase the power of a study, with earlier studies suggesting that as few as 20 matched pairs will adequately detect differences in accelerometry output (Kelly et al. 2005).

In total, 33 participants were recruited for the observational study. Data from two participants were excluded: one participant refused to wear the accelerometers, and a second participant wore the accelerometer belt for a few minutes before removing their belt. The final sample size was 31 (15 males, 16 females, mean (SD) age: 4.4 (0.8) y, height: 104.8 (6.3) cm, weight: 17.7 (2.5) kg, Body mass index (BMI): 16.1 (1.1) kg/m^2). The sample characteristics for the participants involved in the studies in Chapters 4 to 6 are presented in Table 2.1. Initially, only two 7164 Actigraph accelerometers were available for data collection, and therefore fewer participants wore this accelerometer. In the study that compared different generations of accelerometers, (Chapter 7), data were collected from 23 participants who wore the 7164, the GT1M and the RT3 accelerometers. However, the initial analysis revealed significant differences ($p < 0.05$) between accelerometry models with the data from the 23 participants and therefore the study was felt to be sufficiently powered and hence more participants were not sought. The characteristics of the sample in Chapter 7 are presented in Table 2.2.

Table 2.1: Characteristics of the sample (Chapters 4, 5 and 6).

	Mean (SD)		
	All	**Male**	**Females**
No. participants	31	15	16
Age (years)	4.4 (0.8)	4.3 (0.8)	4.3 (0.8)
Height (cm)	104.8 (6.3)	106.2 (5.4)	103.6 (7.0)
Body weight (kg)	17.7 (2.5)	18.0 (2.0)	17.4 (3.0)
Body mass index (kg/m^2)	16.1 (1.1)	15.9 (1.0)	16.1 (1.2)

Table 2.2: Characteristics of the sample (Chapter 7).

	Mean (SD)		
	All	**Male**	**Females**
No. participants	23	10	13
Age (years)	4.5 (0.7)	4.4 (0.7)	4.5 (0.8)
Height (cm)	106.2 (6.2)	108.0 (4.7)	104.8 (7.0)
Body weight (kg)	18.3 (2.5)	18.6 (2.0)	18.1 (2.9)
Body mass index (kg/m^2)	16.2 (1.1)	15.9 (1.1)	16.4 (1.1)

2.5 OUTCOME MEASURES AND INSTRUMENTATION

2.5.1 Direct observation using the CARS

The criterion measure used within the observation study was the Children's Activity Rating Scale (CARS) (Puhl et al. 1990). This is a direct observation scale which has been developed and validated with young children (n = 25, 12 boys, 13 girls, 5 - 6 y) (Puhl et al. 1990). The CARS has been used extensively as a criterion method of measuring physical activity in studies of young children (Finn and Specker 2000; Noland et al. 1990; Oliver et al. 2009; Van Cauwenberghe et al. 2011). Using a scoring system, trained observers can observe and rate children's activity, categorising it into different intensity levels. The scoring system consists of five levels, from sedentary to vigorous intensity activity, that reflect different levels of energy expenditure in children (Puhl et al. 1990). Table 2.3 outlines the CARS levels. Puhl et al. (1990) calibrated the CARS level with twenty-five, 5- to 6-year-old children against indirect calorimetry and heart rate. The significant differences seen in energy expenditure and heart rate between CARS levels, suggested that it could accurately discriminate between different levels of intensity.

Table 2.3: Children's Activity Rating Scale (CARS).

Level 1	Stationary/motionless
Level 2	Stationary/movement of limbs or trunk (very easy)
Level 3	Translocation (slow/easy)
Level 4	Translocation (medium speed/moderate)
Level 5	Translocation (fast or very fast/hard)

(Puhl et al. 1990)

In the current study, Level 1 and 2 activities were collapsed into a 'sedentary' category. While Level 2 activities involve movement of the limbs or trunk, they are stationary activities, therefore in this study Level 2 activities were classified as being sedentary. This approach has also been adopted in earlier studies (De Bock et al. 2010). Levels 3 to 5 involve translocation movement, and refer to light, moderate and vigorous intensity physical activity respectively. Some researchers have interpreted CARS Level 3 activity as being moderate intensity physical activity (Van Cauwenberghe et al. 2011). However, in the original study by Puhl et al. (1990) the MET value for activities classified as being at CARS Level 3 was less than three METs. Therefore, in the current study the Level 3 CARS score was defined as being light intensity physical activity, as activities were less than three times the child's measured MET values. This is in keeping with other studies of pre-school children (De Bock et al. 2010; Sirard et al. 2005). In addition, as discussed in the literature review (Chapter 1, section: 1.3.1, p 2), while three METs has been used as the threshold for moderate intensity activity in adult studies, where the RMR is defined as one MET, equalling 3.5 $ml \cdot kg^{-1} \cdot min^{-1}$ or 4.184 $kJ \cdot kg^{-1} \cdot h^{-1}$ (Ainsworth et al. 2000), children have a higher RMR (Ridley and Olds 2008). It is argued that a higher MET threshold for moderate intensity physical activity is therefore needed for children (Guinhouya and Hubert 2008). Treuth et al. (2004a) proposed 4.6 METs as a threshold for moderate intensity physical activity in children and other authors have argued for six METs for moderate intensity physical activity in children (Guinhouya and Hubert 2008). Finally, Levels 4 to 5 of the CARS were looked at separately as moderate and vigorous intensity activity and were collapsed into a MVPA category as adopted in earlier studies (De Bock et al. 2010; Sirard et al. 2005).

2.5.2 Scoring the CARS

A comprehensive list of activities assigned to the different CARS levels is given in the original article by Puhl et al. (1990) and this is presented in Appendix II g. These categories of activity were used to code the physical activity behaviour of children in this thesis.

The scoring system in the original study involved observing the children's activity and allocating a score between one to five once for each activity level observed and lasting more than 3 s within a 1-minute observation period. This score is then averaged for that 1-minute period. For example, if a child runs fast, sits, then finally walks slowly and each activity lasts more than 3 s, they will score a five, one and three over the minute and an average score for that minute is then calculated as: $\frac{5+1+3}{3} = 3$. This minute of activity would be scored as Level 3 and therefore be classified as light intensity physical activity. At the end of the testing period each child will have a score for the number of minutes they spent at each intensity.

The CARS was designed to be coded over 1-minute observation periods, however, in the current research a modified CARS, as outlined by Sirard et al. (2005) was used, whereby the coding was undertaken over a 15-s period. Given that children engage in short, sporadic bursts of vigorous activity, it has been argued that measuring physical activity over a 1-minute sampling periods may have a 'smoothing' effect, resulting in an underestimation of time spent in vigorous activity (Trost et al. 2005). Therefore, in this research the decision was made to use a 15-s epoch for coding the CARS, to try to reduce the 'smoothing' effect. Using a 15-s epoch was also important to allow cross-comparison with the shorter epoch collected with the accelerometers. A 15-s scoring procedure, as outlined by Sirard et al. (2005), was therefore adopted. However, the other principles outlined in the original study by Puhl et al. (1990) still apply.

2.5.3 Video data collection

To improve the accuracy of the coding system the researcher videoed the children during the data collection period and then retrospectively coded each child using the video data. Two video cameras were set up at each end of the outdoor play area and the researcher used video data from both cameras to code the children's physical activity behaviour. Figure 2.1 shows the view from the first video camera and the position of the second video camera at the CGU5 nursery is noted. Figure 2.2 illustrates the view from the second camera with the digital clock used to retrospectively code the data. One camera was located on a balcony that allowed videoing of the playground area. The second camera was located at the opposite end of the playground on a pillar. The cameras were synchronised with the clock on a laptop PC that was also synchronised with the time on the accelerometers. This video data allowed the researcher to watch the videos, pause and replay to ensure that they coded each 15-s period accurately. In the second nursery, KLN, a similar set up was adopted where cameras were located at opposite ends of an outdoor play area. One camera was positioned on an extended tripod and a second positioned on a roof to capture the playground area on video.

Figure 2.1: Outdoor play area CGU5 nursery, view from camera one.

Figure 2.2: CGU5 nursery, view from camera two.

The video data were coded by the principle researcher, using the CARS (Puyau et al. 2002). Data from both video cameras were necessary to view children in different areas of the playground. There were, however, areas of the playground, such as inside a playhouse and directly under the balcony at CGU5 nursery, where video camera one was located, where the child could not be viewed in either of the video recordings. During the time periods where the child could not be viewed on the video data, the researcher did not code the data, i.e. score the CARS. Only complete 15-s periods of observed activity were included. If a child was out of view within a 15-s period, the entire 15-s period was excluded from the analysis. The corresponding time periods were matched with the accelerometry data, and were also excluded from the analysis.

2.5.4 Reliability of CARS coding

The principal researcher coded the CARS data throughout the research. As there was only one researcher coding the data, the video data allowed for intra-rater reliability checking of the researcher's scoring system and to check for any 'drift' in the coding pattern across the study. Six 10-minute extracts of video data from across the video data collected were re-coded in 2009, one year after the final data collection. The re-coding was undertaken with extracts of video data from six subjects who had originally been coded at different time points of the data collection period (two at the start, middle and end of data collection period). These codes were checked against the original scores allocated in 2007-8.

To evaluate the changes in the mean between the coding sessions, the mean difference (*dm*) and the 95% confidence interval (95% CI) of the mean difference (*dm*), were calculated as plus and minus two standard error (*SE*). Calculations of the confidence interval for the mean difference and standard error are presented in Appendix II h.

To assess reliability, intraclass correlation coefficients (ICC), which are a univariate statistic of relative reliability (Batterham and George 2000), were calculated using the average codes produced for each 15-s epoch. In addition, as the CARS involves ordinal data, where categories are mutually exclusive, specific statistical tests to assess reliability are required and the kappa statistic is commonly used (Sim and Wright 2005). The kappa statistic with 95% CI was also calculated to determine consistency between the ratings. Finally, the percentage agreement (±SD) of the activity categorisation for each 15-s time period was calculated for all time points across the six data samples.

In total, there were 209 pairs of 15-s data (the original and the re-coded data) which were analysed.

The results of the mean difference between the coding sessions are presented in Appendix II h. The mean differences were close to zero (ranging from 0.03 to 0.33) and the confidence intervals included zero, suggesting that there was no systematic bias in the coding.

The results of the ICC analysis are presented in Appendix II h. The ICC values ranged from 0.72 to 0.96. While there is no generally agreed 'cut-off', categories for interpreting ICC values, the closer the ICC is to one, the better the reliability. Fleiss (1986) recommends that ICC values between 0.4 and 0.75 indicate 'fair to good' and values > 0.75 suggest 'excellent' relative reliability. Five of the ICC values in this study were > 0.75 and one value was close to this value, suggesting 'good' to 'excellent' reliability. It should be noted that while the confidence interval for subject two ranged from 0.43 up to 0.86, the ICC value of 0.43 is still considered to be 'fair' reliability according to the Fleiss (1986) classification.

The intra-rater reliability using the kappa was = 0.71 (p <0.0001), 95% CI (-0.07, 1.49). The average percentage agreement (SD) of 15 s physical activity categorisation across time points was 74.6 % (9.2).

One means of interpreting the kappa value is to use the scale outlined by Landis and Koch, (1977), whereby the closer the value is to 1 the better the reliability. In the scale, values ≤ 0 = poor, 0.01 - 0.20 = slight, 0.21 - 0.40 = fair, 0.41 - 0.60 = moderate, 0.61 - 0.80 = substantial and 0.81 - 1 = almost perfect reliability. Given that the kappa value in this study is 0.71 which is within the range 0.61 - 0.81, this can be interpreted as 'substantial agreement'.

These results suggest that there was good intra-rater reliability of the coding used across the study. This is in keeping with earlier studies where both intra and inter-rater reliability of the CARS coding have been found to be high (DuRant et al. 1993; Puhl et al. 1990; Sirard et al. 2005).

2.5.5 Accelerometry equipment

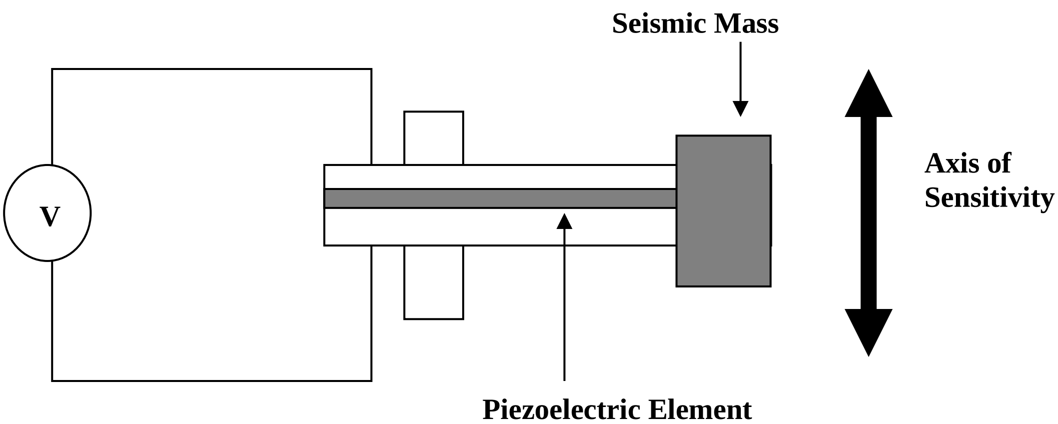
The 1-hour direct observation data collection reported in Chapters 4 to 7 used the GT1M Actigraph uniaxial accelerometer, the 7164 uniaxial accelerometer and the RT3 triaxial accelerometer. In recent years, Actigraph have released the GT3X and GT3X+ triaxial accelerometers. These monitors became commercially available in 2009 and to date there are few published studies with pre-school children available, and of those published, the GT3X/GT3X+ accelerometers have been used in a uniaxial setting (Kahan et al. 2013). The 1-hour direct observation data collected in this thesis was undertaken in 2007 and 2008 and therefore pre-dates the availability of the GT3X/GT3X+ model. The final study reported in Chapter 8 was undertaken in 2010 and used the GT3X model, in a uniaxial mode. The technical specifications of the GT1M, GT3X, 7164 and RT3 accelerometers will be discussed in the following section.

2.5.5.1.1 7164 Accelerometer

The 7164 accelerometer (Figure 2.3) is a uniaxial accelerometer which, prior to 2005, was widely used in physical activity research. The 7164 detects acceleration in the range of 0.05 to 2.0 *G*, in the vertical plane, by the means of a piezoelectric acceleration sensor. This sensor consists of a piezoelectric element and a seismic mass which is attached to the free end of a piezoceramic cantilever arm (Figure 2.4) (Chen and Bassett 2005). The opposite end is mounted on the monitor's electronic circuit board (John and Freedson 2012). Acceleration forces act on the seismic mass, which causes the piezoelectric cantilevered arm to bend and a charge proportional to the strain on the piezoelectric element is generated (Tryon and Williams 1996). The electric charge which is generated is then filtered by an analogue band-pass filter and digitalised by an 8-bit analogue-to-digital converter (ADC) at 10 samples per second (Manufacturing Technology Inc. 2001). This signal also undergoes ‘full-wave rectification’, which means it is transformed to absolute acceleration values (John and Freedson 2012). The ADC value is then summed over a predefined period of time known as an epoch and the output is reported in activity counts per epoch. The epoch can be set at anything between 1 s and several minutes.

Figure 2.3: The 7164 Actigraph accelerometer.

Figure 2.4: Schematic representation of cantilever beam in the 7164 accelerometer.



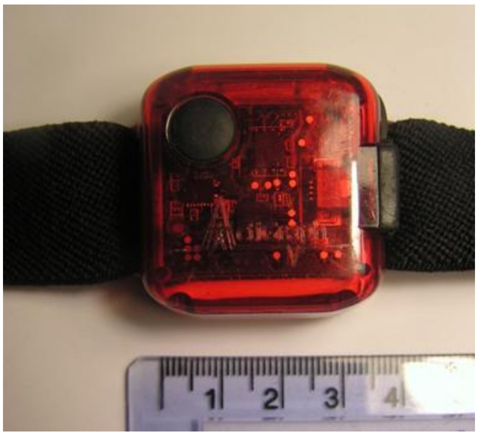
Adapted from Chen & Bassett Jr (2005)

The acceleration signal is band limited between 0.25 to 2.5 Hz, which, according to the manufacturers, has been selected to detect human movement and exclude motion from other sources (Manufacturing Technology Inc. 2001). According to Chen and Bassett Jr (2005), a limitation of the piezoelectric accelerometer is that they have a phenomenon known as 'leakage', in which the initial change in charge dissipates over time even if the static loading is still present. This means that only dynamic events can be reliably monitored and they are not well suited to measuring different angles with respect to gravity, such as different body postures.

2.5.5.1.2 GT1M and GT3X accelerometers

The internal technology of the GT1M and GT3X is different to the earlier 7164 model. Both the GT1M and GT3X consist of a solid-state accelerometer using an integrated micromachined monolithic integrated circuit chip (polysilicon) to detect acceleration (Analog Devices 2007). The GT1M uses a dual-axis micromechanical system accelerometer and the GT3X makes use of a triaxial capacitive micromechanical system (John and Freedson 2012). The sensor is suspended by springs over the surface of silicon water and provides a resistance against acceleration forces (Analog Devices 2007). One of the main differences with the accelerometer sensors in the GT1M and GT3X from the 7164 is that they can detect both dynamic acceleration (e.g. as a result of motion) and static acceleration (e.g. as a result of gravity forces). This function allows the GT3 to have a inclinometer, so it can be used to detect position of the body, although the validity of the inclinometer to accurately detect body position is still under investigation and has not yet been validated.

Figure 2.5: GT1M Actigraph accelerometer.

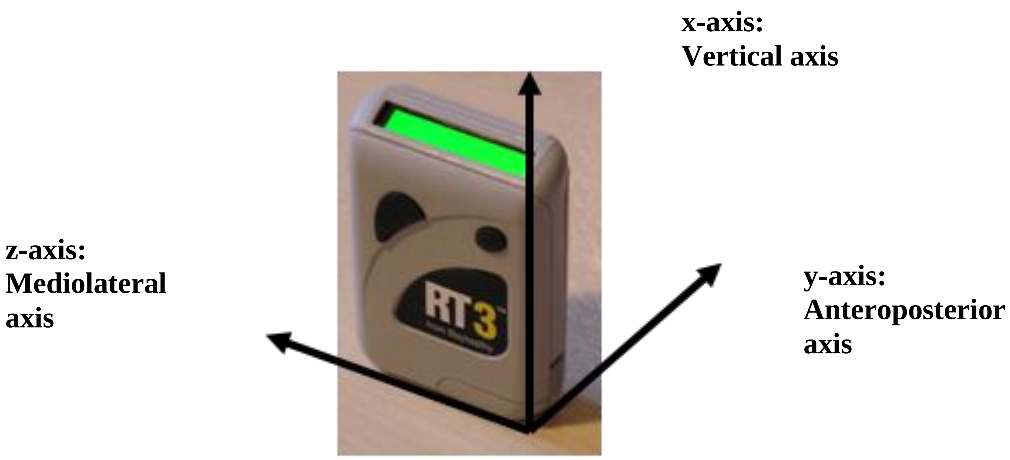


The GT1M and GT3X accelerometers use a 12-bit analogue-to-digital converter digitalised at 30 Hz, and similar to the 7164, the GT1M and GT3X record acceleration in the range 0.05 to 2.0 *G*, making use of a band-pass filter that excludes signals outside the range 0.25 to 2.5 Hz. In addition, the GT1M and GT3X have an option to provide output in a pre-filtered mode and firmware released in 2010 allows data on the *G* force, sampled every 0.033 s, to be collected. The GT1M and GT3X are exactly the same in all specifications, other than the fact that the GT3X has a triaxial accelerometry sensor, while the GT1M has a uniaxial accelerometry sensor. Studies suggest that the GT1M and GT3X have good inter-model reliability, with similar output between reported between models when used in a uniaxial setting (John et al. 2010; Kahan et al. 2013; Kaminsky and Ozemek 2012).

2.5.5.1.3 RT3 accelerometer

The RT3 is a triaxial accelerometer which measures acceleration in three orthogonal planes: vertical, anteroposterior, mediolateral planes. Figure 2.6 illustrates the RT3 with its sensitive axes.

Figure 2.6: RT3 accelerometer and axes of measurement.



Similar to the Actigraph models, acceleration is measured and converted to a digital signal within the range 0.05 to 2.0 *G* and is sensitive to the range 2 - 10 Hz (Powell and Rowlands 2004). The RT3 has been reported to have good reliability during locomotive activity (CV < 6%) (Powell and Rowlands 2004) and reasonable validity in treadmill walking with adults (n = 25, 18 – 65 y) (Hendrick et al. 2010). There is less information available on the specifications of the RT3, such as the analogue-to-digital converter and the frequency of digital converter.

Table 2.4 summarises the technical specification for the accelerometers used in the current thesis.

Table 2.4: Technical specification of accelerometers used in this thesis.

	7164	GT1M	GT3X	RT3
Size (cm)	5.1 x 4.1 x 1.5	3.8 x 3.7 x 1.8	4.6 x 3.3 x 1.5	7.1 x 5.6 x 2.8
Weight (g)	42.5	27	27	65.2
Sensitive Axes	Vertical (x)	Vertical (x)	Vertical (x); ML (z); AP (y)	Vertical (x); ML (z); AP (y)
Sensitive range (*G*)	0.05 - 2.0	0.05 - 2.0	0.05 - 2.0	0.05 - 2.0
ADC	8-bit	12-bit	12-bit	NA
Frequency of digital converter	10 Hz	33.33 ms (30 Hz)	33.33 ms (30 Hz)	NA
Band-pass filter	0.25 to 2.5 Hz	0.25 to 2.5 Hz	0.25 to 2.5 Hz	2 to 10 Hz
Battery	2430 coin cell lithium battery	Prismatic lithium ion battery	Prismatic lithium ion battery	2 x AAA batteries
Memory	64 KB	1 MB	16 MB	NA
Data storage capacity	1 min epoch: 22 days 1-s epoch: 8.9 hour	1-min epoch: 365 days 1-s epoch: 4 days	1-min epoch: 5643.4 days 1-s epoch: 94 days	1-min epoch: 7 - 21 days 1-s epoch: 3 - 9 hours

ADC: analogue-to-digital converter; AP: anteroposterior; ML: mediolateral; NA: not available.

2.5.6 Mechanical calibration of accelerometers

Calibration of the accelerometers used in this study was undertaken using a mechanical set up based on a calibration unit developed by Brage et al. (2003). This allowed for 'unit calibration' to ascertain whether accelerometers were providing similar information under test conditions (Welk 2005). Determining the inter- and intra-unit reliability of accelerometers is important as the accelerometers were used over an extended period of time. Calibration of accelerometers was undertaken both pre- and post-field data collection. The purpose of the calibration was to determine any sources of variation both pre- and post-field testing and to identify any technical problems with the accelerometers.

The details of the calibration studies are reported in Appendix II i. The results suggested that the GT1M models had good intra-unit reliability, with coefficient of variation (CV%) values ranging from 0.66 pre-field data collection and 1.57 post-field data collection. The GT1M also had good inter-unit reliability between accelerometers of the same model with ICCs that were very close to 1 (perfect agreement) (0.99). The RT3 and 7164 both had higher CV% values for intra-unit reliability (RT3 model: 5.19 pre- and 5.93 post-; 7164 model: 8.28 pre- and 4.75 post-field data collection). These findings are similar to the values reported in other mechanical calibration studies where the GT1M has been reported as having better reliability than the RT3 and the 7164 accelerometer (Esliger and Tremblay 2006; Rothney et al. 2008). The CV% values of the accelerometer models and the ICC values were similar both pre-and post-data collection.

Two of the 7164 accelerometry units failed during calibration testing pre-field data collection (data output was 0) and both were excluded from the analysis and not used in the field data collection. Six 7164 accelerometers were included in the pre- and post-field data collection tests. One GT1M was found to be faulty prior to data collection (it did not charge) and was returned to the manufacturer. Nine GT1M accelerometers were tested in the pre- and post-field data collection. During the field data collection one RT3 accelerometer stopped charging and had to be excluded from the research. Ten RT3s were tested pre-field data collection and nine were tested post-field data collection.

During the calibration process it was noted that the GT1M accelerometers were not synchronising with the PC clock. This problem was not observed with the other models (7164 and the RT3) during simultaneous mechanical calibration. The GT1M accelerometers were found to be 15 – 20 s out of sync from the PC clock. Appendix II m illustrates some of the raw data in 1-s epochs for the GT1M and the 7164 accelerometers collected during mechanical calibration, where the start and stop times with the PC clock are highlighted. It can be seen that the 7164 accelerometers are synchronised with the start and finish times of the mechanical calibration unit (which was timed with the PC clock). The GT1M accelerometers were 15 – 20 s out of sync from the PC clock.

2.6 PROCEDURES AND ANTHROPOMETRIC MEASURES

Following recruitment, anthropometric measures of the children's height and weight were taken to allow characterisation of the sample population. Weight was measured to the nearest 0.1 kg using Seca analogue scales (Seca United Kingdom, Birmingham, UK) and height to the nearest 0.1 cm using a Seca Leicester portable stadiometer. To ensure standardisation in the measurement of height and weight the protocol outlined in the National Diet, Nutrition and Dental Survey of Children aged 1 ½ to 4 ½ years, 1992 – 3, was followed (Department of Health, 1995).

To measure the child's height, the child was asked to stand barefoot with their heels touching the back of the stadiometer. The child was asked to look straight ahead with arms relaxed by their sides. The researcher gently held the child's head in two hands so that light upwards pressure was applied under the jaw anteriorly and occiput (base of the skull) posteriorly to provide maximum extension of the spine. Care was taken not to tilt the head and to maintain the 'Frankfurt' position of the head, whereby the inferior aspect of the orbit was parallel with upper margin of the ear canal (Office of Population Censuses and Surveys. Social Survey Division 1995). The child was asked to breathe in and then out and to relax their shoulders without lifting their heels from the ground. The horizontal head plate was then lowered until it made contact with the highest point of the child's head and height was recorded to the nearest 0.1 cm.

Children were weighed in light clothing and were asked to stand barefoot in the centre of the Seca scales with arm by their sides. Weight was measured to the nearest 0.1 kg.

Using the height and weight data, the children's BMIs were calculated. BMI was calculated as the body mass in kg divided by the square of the height in metres (kg/m^2). In addition, the BMI scores were expressed as an age and sex adjusted standard deviation score (*z*-score), relative to the UK 1990 population reference data (Cole et al. 1995). The proportion of children classified as being 'healthy weight', 'overweight' and 'obese' was determined using the international cut off points for BMI for overweight and obesity by sex and age (Cole 2002).

Prior to data collection, the GT1M, 7164 and RT3 accelerometers were set up to collect data in 1-s epochs. The accelerometers were initialised using a laptop PC and were set to start collecting the data at a pre-determined time (usually 1 hour prior to data collection commencing). In this way the time on the PC was synchronised with the internal accelerometer clock. The digital clock for the video camera was also synchronised with the time on the PC. The video cameras were set to record data 15 minutes prior to children entering the outdoor play area.

The GT1M and the 7164 were positioned on an elastic belt worn around the child's waist, over the mid axillary line positioned over the child's hip. Attachment of accelerometers close to the centre of body mass has been advocated as the optimal position (Puyau et al. 2002). While wrist-worn accelerometers are available, Trost et al. (1998) argues that wearing accelerometers on the wrist may increase subject reactivity, increase the risk of the subject tampering with the device, and could pick up extraneous arm movements. For these reasons the hip position for the GT1M and the 7164 accelerometers was chosen. One accelerometer was positioned anterior to the other and the order of the positioning was randomised (Figure 2.7).

The RT3 was clipped onto the child's waist band of their trousers or skirt. During piloting it was observed that the additional weight of the RT3 meant that when it was positioned on the elastic belt it caused extra movements of the belt and subsequently extra movements of all the accelerometers. For this reason it was necessary to secure the RT3 to the waist band of the child's clothing. It was not possible to collect data from children who did not wear trousers or skirt, e.g. if they were wearing a summer dress. On these occasions the data were collected at a later date. The 7164, GT1M and the RT3 were worn on opposite sides to prevent the monitors coming into contact with each other, and the position was randomised.

Figure 2.7: Positioning of accelerometers.

Following initialisation of the accelerometers the children were fitted with their accelerometers by the researcher. The nursery teachers took the children into the playground for their usual, time-tabled, outdoor playtime of approximately 1 hour. The children were free to run and play and no structure was put in place for this session. If a child removed their accelerometer during the period of data collection, or tampered with the accelerometer they were wearing then these data were excluded and not used in the final analysis.

Data were collected from CGU5 in August 2007 and from KLN between March and May 2008. During data collection only three to four children were wearing accelerometers at a time, therefore data were collected from the sample over several sessions, with six, 1 hour sessions being undertaken at the CGU5 and four, 1 hour sessions being undertaken at KLN. Sixteen participants were recruited from KLN (7 males, 9 females, mean (SD) age: 4.8 y, height: 108.4 (4.8) cm, weight: 19.3 (2.2) kg, BMI: 16.4 (1.2) kg/m^2) and 15 participants were recruited from the CGU5 nursery (8 males, 7 females, mean (SD) age: 3.8 (0.7) y, height: 101.1 (5.5) cm, weight: 16.0 (1.7) kg, BMI: 15.6 (0.8) kg/m^2). The presence of younger children at CGU5 is explained by the data collection being undertaken at the start of the children's academic year, whereas at KLN, children were older as data were collected towards the end of the pre-school year.

2.7 PILOT AND FAMILIARISATION

The procedures were piloted over two sessions at CGU5 in July 2007 and one session at KLN nursery in February 2008. The pilot sessions were undertaken to check the procedures used in data collection and to determine the optimal set-up for the video cameras to capture the playground area. Initially one camera was used to collect data, but it became apparent that two were required to capture children in the different areas of the outdoor play areas. The pilot also allowed the children to become familiar with the researcher and to allow them to adjust to wearing the accelerometers. Although earlier studies suggest low reactivity from wearing activity monitors (Vincent and Pangrazi 2002), there appeared to be some initial changes to behaviour in the children wearing the accelerometers. However, the novelty of wearing the accelerometers appeared to diminish within the pilot session. Following the pilot session, a change to the position of the RT3 was undertaken as outlined earlier. In addition, the accelerometry belts were shortened to improve comfort for the children.

2.8 DATA PROCESSING AND ANALYSIS

Following data collection the data from the accelerometers were transferred to Excel for processing. Using a macro created in Visual Basic for Applications, the data collected in 1-s epochs were re-integrated into 3-, 5-, 15-, 30- and 60-s epochs. Cut-point thresholds were applied to the data to allow for calculation of time spent at each of the intensities (light, moderate and vigorous intensity physical activity), as well as time spent in sedentary behaviour. Details of the thresholds applied are given within each of the study chapters. Data were processed using SPSS for Windows (version 15.0 and version 17.0). Normality plots and analysis were conducted using Shapiro-Wilks as the sample size was less than 50, except in Chapter 8 where sample sizes were greater than 50 and the Kolmogorov-Smirnov statistic was used.

Where the data were normally distributed ($p > 0.05$) the means and standard deviations (SD) were presented. Bar charts with the error bars to indicate the SD were used to present the normally distributed data graphically. Parametric inferential statistics were used to determine differences or explore associations (details are provided in the respective study chapters). Where the data were found to be not normally distributed ($p < 0.05$) the median and interquartile ranges (IQR) were presented. With non-parametric data, the data were presented graphically using box plot graphs presenting the median, IQR and minimum and maximum values. Non-parametric inferential statistics were used to determine differences and test for associations (details are given within the respective study chapters).

To evaluate the agreement between measures it is recognised that while correlation analysis can measure the strength of a relationship it can overlook systematic differences between variables (Oliver et al. 2007b). Therefore, Bland and Altman plots with 95% (± 1.96 SD) limits of agreement (LOA) were determined (Bland and Altman 1986).

Further details of the specific analysis undertaken in each of the studies are discussed within the respective study chapters.

CHAPTER 3 : EXAMINATION OF EPOCH EFFECT ON ESTIMATES OF PHYSICAL ACTIVITY OF PRE-SCHOOL CHILDREN

3.1 INTRODUCTION

The content of this chapter was published as part of a review paper in 2008 (Reilly et al. 2008). The findings from the publication are presented in this chapter. At the time of publication, no studies had previously explored the implications of epoch on estimates of physical activity in pre-school children. However, since 2008 there are now small number of studies published in this area. The implications of these studies will be considered in Chapter 4 which explores the accuracy of different epochs. In addition, changes to the physical activity recommendations for health for pre-school children were published in the UK in 2011. The findings of this first epoch chapter will also be considered in greater depth in relation to the new physical activity recommendations in the general discussion chapter (Chapter 9).

When using accelerometers it is necessary to pre-set the sampling intervals or epochs for data collection, which conventionally have been between 15 s and 60 s. Early accelerometry studies tended to use 1-minute epochs, in part due to the memory limitations of earlier models of accelerometers, which did not have the capacity to collect data over extended periods of time at shorter epochs (Rowlands et al. 2006). It has been argued that use of longer epochs (e.g. 1 min) may not accurately capture the short, sporadic bursts of vigorous activity thought to be typical of young children (Trost 2001). As a result of the perception that children tend to engage in intermittent bouts of high intensity physical activity, it is widely believed that shorter epochs would be more appropriate to use in studies with children. This perception is based on the frequently cited study in which 15, 6- to 10-year-old children were observed over 3 days (Bailey et al. 1995). More recent studies using direct observation and heart-rate monitoring to measure patterns of physical activity in children suggest a much more sedentary pattern of behaviour with limited physical activity and patterns of physical activity much more like adults (Cardon and De Bourdeaudhuij 2007; McKee et al. 2005). Despite this, the concern with using longer epochs is that the short burst of high-intensity activity will be averaged within the epoch which includes longer periods of low intensity activity. As a result, high intensity physical activity could be misclassified within that epoch, leading to an underestimation of true levels of high intensity physical activity (Trost et al. 2005).

Two studies have investigated the influence that epoch can have on time spent at different intensities of physical activity in school-aged children. Rowlands et al. (2006) compared outcomes with 1- and 60-s epochs in RT3 accelerometry data collected from 25, seven-to-11 year olds over 1 hour. The authors reported that the differences between the two epochs were minimal and the main effect was that some of the 'very hard' intensity physical activity was misclassified as 'hard' intensity physical activity when the longer epochs were applied. Nilsson et al. (2002) found no significant difference in the amount of time spent in light and moderate intensity physical activity with the Actigraph in 16, seven-year-old children when using different epoch settings (5-, 10-, 20-, 40- and 60-s epochs). However, some of the vigorous intensity physical activity was misclassified as moderate intensity physical activity with the longer epochs. One practical solution to this potential problem when using longer epochs is to classify moderate and vigorous activity together, as MVPA (Reilly et al. 2004). This is also biologically and clinically meaningful as prior to 2009 the public-health targets for physical activity in children and adolescents were expressed in terms of MVPA. Although the more recent recommendations for health are for pre-school children to engage in 180 min of total physical activity (TPA) per day (Department of Health, Physical Activity, Health Improvement and Protection 2011), as discussed in the background to the current thesis (Chapter 1, p. 8), the relationship between different intensities of physical activity and health is not clear. Therefore, to gain further understanding of the relationship, accurate measurement of the different dimensions of physical activity is important.

The implications of different epochs on estimates of physical activity and sedentary behaviour have not been investigated with pre-school children. In addition, the question regarding whether shorter epochs offer a more accurate means of quantifying physical activity levels in pre-school children has not been explored (Welk et al. 2000b).

The aim of this study was:

- To examine the influence of epochs on estimates of MVPA and sedentary behaviour of pre-school children with 7 - 10 days of free-living accelerometry data.

3.2 METHODS

A secondary data analysis of habitual accelerometry data collected from 32 children (5 and 6 years) over 7 - 10 days was undertaken. Data had been collected as part of a larger study conducted in 2002 into physical activity levels of pre-school children in Glasgow

(Montgomery et al. 2004). All parents had given written informed consent to participation and the study had approval from Yorkhill Hospital Research Ethics Committee. Table 3.1 outlines the sample characteristics. In this sample 69% were classified as 'healthy weight' and 31% were classified as overweight/obese, i.e. with a BMI at or above the 85th centile relative to UK population reference data (Cole 2002).

Table 3.1: Characteristics of sample (Chapter 3).

	Mean (SD)		
	All	**Male**	**Females**
No. participants	32	17	15
Age (years)	5.9 (0.7)	6.1(0.6)	5.7 (0.8)
Height (cm)	115.1 (6.4)	120 (5.6)	110 (7.2)
Body weight (kg)	21.7 (4.2)	22.6 (4.3)	20.9 (3.9)
Body mass index (kg/m^2)	16.3 (1.8)	16.5 (1.9)	16.1 (1.7)

3.2.1 Equipment

Data were collected using 7164 Actigraph accelerometers set to collect data in 15-s epochs. The 7164 has been used extensively in physical activity research with young children (Fisher et al. 2005b; Jackson et al. 2003; Kelly et al. 2007) and has been found to be a valid and reliable instrument (de Vries et al. 2006; de Vries et al. 2009).

3.2.2 Procedure

Children wore the 7164 accelerometer during waking hours over a 10-day period (mean (SD) days: 9.3 (1.3)). The mean registered time recorded per day was 654 min (approximately 11 hour/day) which is considered to be sufficient hours per day to constitute a valid day (Anderson et al. 2005).

3.2.3 Data Analysis

The 15-s epoch accelerometry data were transferred to Excel and reintegrated into 30- and 60-s epochs using the method adopted by Nilsson et al. (2002). To estimate time spent at different intensities it was necessary to convert the raw accelerometry counts using previously published cut-points. This was achieved by applying published cut-off values for accelerometry output to convert this to time spent in sedentary behaviour, moderate and vigorous intensity physical activity. Using a programme developed with Visual Basic for Applications, the data were processed applying the cut-points defined by Puyau et al. (2002) to classify the data into time spent in moderate and vigorous intensity physical activity as well as grouping these into MVPA. In addition, cut-points for sedentary behaviour by Reilly et al. (2003) were applied. These cut-points, which were developed for 1-min epochs, were divided for the respective epochs. While other cut-points have been developed specifically for the pre-school population many of these predate the publication of this study (Van Cauwenberghe et al. 2011) and at the time of drafting the publication the decision was made to use the Puyau et al. (2002) as they had been validated for use in children using indirect calorimetry. Which cut-point is most accurate will be considered in Chapter 6.

Table 3.2 presents the cut-points applied to the accelerometry data for the different epochs.

Table 3.2: Cut-points applied.

		Epochs (seconds)	
	15 s	**30 s**	**60 s**
Sed[1] (counts/epoch)	< 275	< 550	< 1100
MPA[2] (counts/epoch)	≥ 800 to < 2050	≥ 1600 to < 4101	≥ 3200 to <8200
VPA[2] (counts/epoch)	≥ 2050	≥ 4100	≥ 8200
MVPA[2] (counts/epoch)	≥ 800	≥ 1600	≥ 3200

MPA: moderate physical activity; MVPA: moderate-to-vigorous physical activity;
[1]*Reilly et al.* (2003) *cut-points;* [2]*Puyau et al.* (2002) *cut-points*; *Sed: sedentary behaviour; VPA: vigorous physical activity.*

The number of minutes spent within each intensity was calculated. The data were then analysed within SPSS (version 15.0). The data were checked for normality using the Shapiro-Wilks test as the sample size was less than 30. Results of the normality testing are presented in Appendix III (Appendix Table III.i). Data were found to be significantly different from the normal distribution (p < 0.05) and therefore non-parametric statistics were used to analyse the results.

Descriptive statistics on the median and interquartile (IQR) ranges were reported and the non-parametric test, the Kruskall-Wallis one-way ANOVA, was applied to test for any significant epoch effect (between 15-, 30- and 60-s epochs). Post-hoc analysis was undertaken using the Mann-Whitney test, the non-parametric equivalent of the independent *t*-test (Field 2012), to test for differences between the different epoch settings. To protect against a type I error, the Bonferroni correction was undertaken (α/number of comparisons) (Field 2012) (e.g. 0.05/4 = 0.0125). Significance level was therefore $p < 0.0125$.

3.3 RESULTS

The median and IQR were explored and are presented in Table 3.3.

Table 3.3: Median (IQR) minutes per day spent at each intensity.

		Epoch (seconds)		
		15 s	**30 s**	**60 s**
Sed[1] (min/day)	**Median**	532.6	531.5	529.5
	IQR	(140)	(148)	(145)
MPA[2] (min/day)	**Median**	25.9	22.5	16.0
	IQR	(21.3)	(20)	(18)
VPA[2] (min/day)	**Median**	1.3	0.5	0.0
	IQR	(3)	(1.5)	(1.0)
MVPA[2] (min/day)	**Median**	28.0	23.5	17.0
	IQR	(22.6)	(20.6)	(18.0)

MPA: moderate physical activity; MVPA: moderate-to-vigorous physical activity;
[1]*Reilly et al.* (2003) *cut-points;* [2]*Puyau et al.* (2002) *cut-points*; *Sed: sedentary behaviour; VPA: vigorous physical activity.*

It can be seen from Table 3.3 that there was little difference in the median number of minutes spent in sedentary behaviour (approximately 3 minutes over a day between 15- and 60-s epochs), however, for MVPA there is a difference of 11 minutes between the median values for 15- and the 60-s epochs. Figure 3.1 presents the median and IQR of the minutes per day spent in sedentary behaviour and Figure 3.2 for time spent in MVPA per day for different epoch lengths (15, 30 and 60 s).

Figure 3.1: Box plot of median (IQR) minutes per day spent in sedentary behaviour.

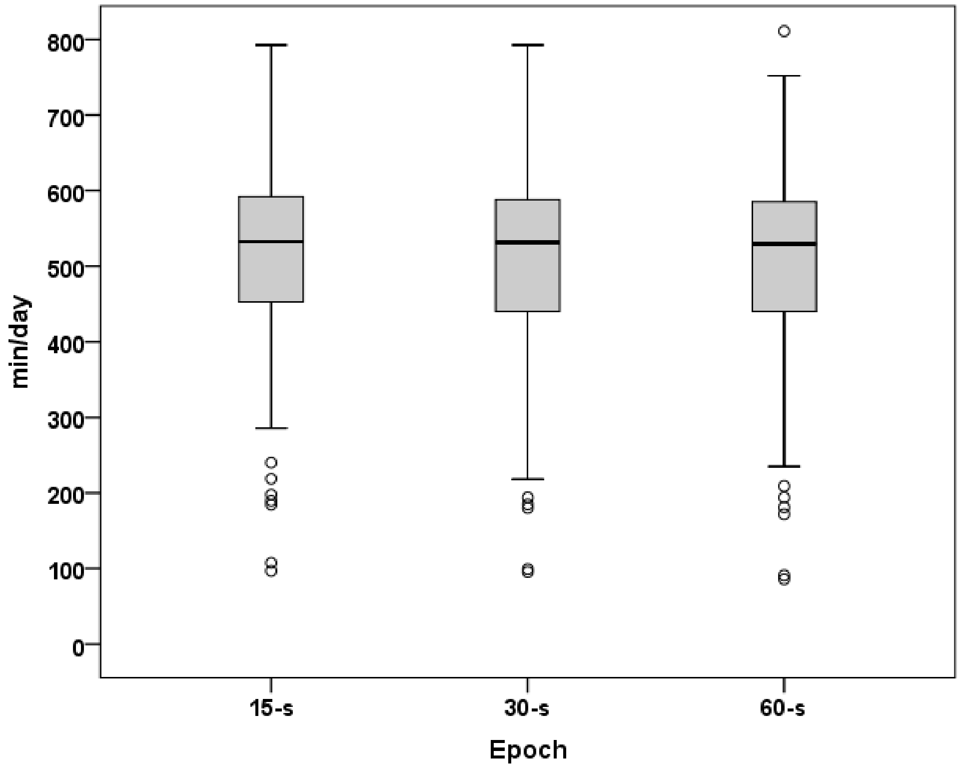
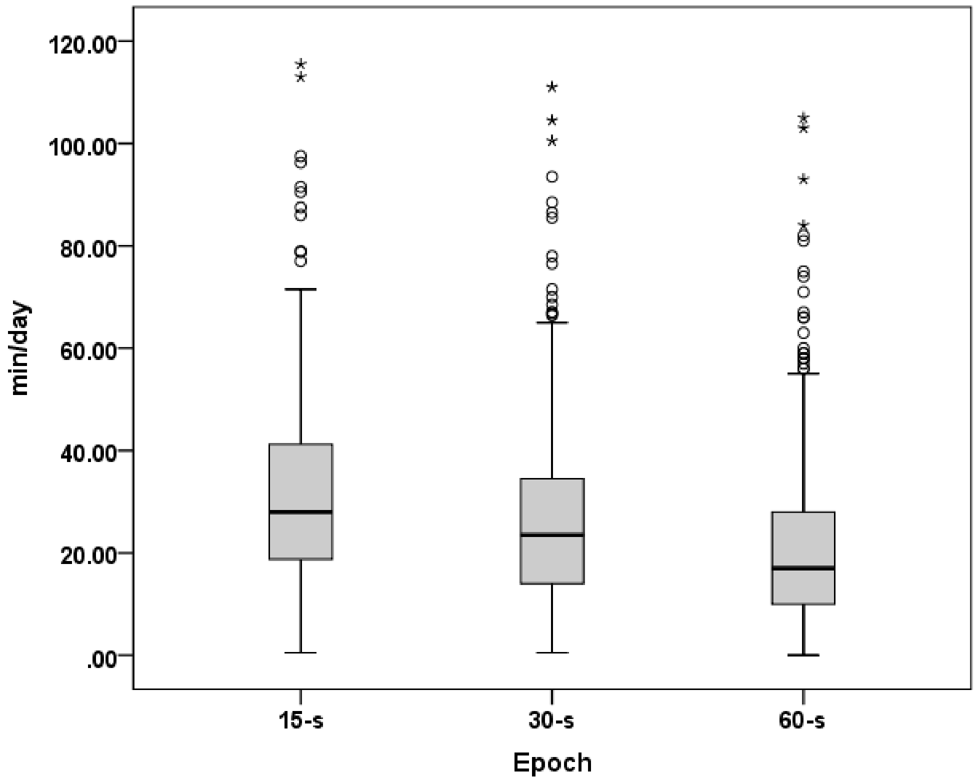


Figure 3.2: Box plot of median (IQR) minutes per day spent in moderate-to-vigorous physical activity.



The results of the MVPA box-plot illustrate the distribution of the data, highlighting that there were several outlier participants who had higher levels of MVPA, and similarly several outlier participants spent less time in sedentary behaviour. Results of the Kruskall-Wallis test revealed that there was a significant difference between all epochs for moderate intensity physical activity (H(2) 57.56, $p < 0.05$), vigorous intensity physical activity (H(2) 149.41, $p < 0.05$) and MVPA (H (2) 63.5, $p < 0.05$). There was no significant difference for epochs for sedentary behaviour (H(2) 0.43, $p = 0.81$).

Post-hoc analysis was undertaken using the Mann-Whitney test to explore differences between the epochs for time spent engaged in MVPA. Effect sizes (*r*) were calculated using the *z*-scores divided by the square root of the number of observations (Equation 3.1). It was found that there was a significant difference between 15- and 30-s epochs and between 15- and 60-s epochs (Table 3.5)

Equation 3.1: Calculation of effect size

$$\frac{z - \text{score}}{\sqrt{\text{number of comparisons}}}$$

(Field 2012)

The *r* effect sizes range from 0, suggesting no effect to 1 suggesting a perfect effect. Table 3.4 presents the effect size criteria proposed by Cohen (1988; 1992) for small and large effect sizes which are widely accepted (Field 2012).

Table 3.4: Interpretation of effect sizes

r	***Categorisation of effect***	***Explanation of effect***
$r = 0.10$	Small effect	explains 1 % of the variance
$r = 0.30$	Medium effect	explains 9 % of the variance
$r = 0.50$	Large effect	explains 25% of the variance

(Adapted from Field 2012)

It can be seen in Table 3.5 that the effect sizes were small between 15- and 30-s epochs, and there was a medium effect size between 15- and 60-s epochs.

Table 3.5: Effect sizes for MVPA activity

Epoch difference	*p*-value	*z*-scores	Effect size (*r*)
15 s & 30 s	0.001	-3.82	-0.16
15 s & 60 s	0.001	-7.84	-0.32

3.4 DISCUSSION

The results suggest that there was a significant epoch effect between the different epochs used for time spent in MVPA, in particular between 15- and 60-s epochs. This could suggest that if data are recorded using 60-s epochs, the recorded number of minutes of MVPA during a day could be underestimated. However, the difference for time spent in MVPA between 15 s and 60 s was small (11 min/day), and there was a moderate effect size. It is unclear whether this difference is biologically or clinically significant. This study provides some evidence to support the widespread perception that shorter epochs are essential to accurately measure physical activity in young children.

It is notable that in this study pre-school children spent a limited amount of time per day engaging in vigorous intensity physical activity and the median minutes ranged from 0 to 1.3 min/day (at 60- and 15-s epochs respectively). There was a much larger epoch effect for time spent in moderate intensity physical activity. In earlier studies the epoch effect was seen with high intensity physical activity (Nilsson et al. 2002; Rowlands et al. 2006) and this does not seem to be supported in the findings from the current study. Although not presented in the Reilly et al. (2008) review paper, the time spent in LPA using the Puyau et al. (2002) cut-points (≥ 200cpm to 800 counts/15 s), would result in a median (IQR) minutes of 144.9 (56.3) min at 15-s epoch and 173.5 (79) min at 60-s epoch. This difference of 28.6 min/day could have considerable implications for estimates of a pre-school child's TPA.

In conclusion, an epoch effect was observed, with the 15-s epoch resulting in significantly ($p < 0.05$) more time in per day in MVPA than 60-s epochs. There was a moderate effect size ($r = -0.32$) for this difference of 11 min/day.

CHAPTER 4 : AN INVESTIGATION INTO WHICH EPOCH IS MOST ACCURATE FOR MEASURING PHYSICAL ACTIVITY OF PRE-SCHOOL CHILDREN

4.1 INTRODUCTION

This chapter is based on the researcher's publication, Hislop et al (2012b), published in Pediatric Exercise Science in 2012.

Since publication of the Reilly et al. (2008) review paper in 2008, a number of studies have explored the effect of using different epochs on estimates of physical activity under free-living conditions in school-aged children (Edwardson and Gorely 2010; McClain et al. 2008) and pre-school-aged children (Mahar et al. 2008; Ojiambo et al. 2011; Reilly et al. 2008; Vale et al. 2009). However, few studies have examined the accuracy of shorter epochs (i.e. have compared the estimates of physical activity from different epochs against a criterion method). Also the use of 60-s epochs for the assessment of high intensity activity continues to be questioned (Nilsson et al. 2002; Trost et al. 2005).

Edwardson and Gorley (2010) collected accelerometry data over 7 days from 311, seven- to eleven-year-old children. The authors reported that when estimating MVPA, 5- and 60-s epochs were not comparable and 5-s epochs resulted in significantly more minutes of MVPA than 60-s epochs ($p < 0.01$). This epoch effect has been supported in studies of younger children. Vale et al. (2009) found a significant difference in time spent in MVPA ($p < 0.001$) in 5- and 60-s epochs for accelerometry data collected from young children over 4 school days (28 males, 31 females, mean (SD) age: 4.3 (1.1) y). A mean difference of approximately 17 minutes of time spent in MVPA was reported between the two epochs used. In contrast Reilly et al. (2008) reported a median difference of 11 min/day when 15- and 60-s epochs were compared in free-living data collected from 32, five and six year olds over 7-10 days, concluding that the biological significance of the differences was not clear (Reilly et al. 2008).

In a study involving 72 pre-school children (mean (SD) age: 3.9 (0.6) y), Mahar et al. (2008) reported that there were significant differences for estimates of time spent at moderate, vigorous and MVPA when 1-s epochs were reintegrated into 3-, 5-, 15-, 30- and 60-s epochs. However, in an earlier study Nilsson et al. (2002) reported that there were no differences in time spent in moderate activity but that there were significant differences in estimates of time spent in 'hard' and 'very hard' intensities when 5-s epochs were reintegrated into 10-, 20-, 40- and 60-s epochs with data collected from 16, seven-year-old children over 4 days. Other studies have supported the fact that longer epochs can lead to misclassification of high intensity or vigorous activity (Dorsey et al. 2009; McGrath and Hinckson 2009; Rowlands et al. 2006).

A limitation of many of the studies to date is the lack of a criterion measure of intensity to determine which epoch provides the most accurate estimate of physical activity intensity. To date, only one study by McClain et al. (2008) has evaluated accelerometry epoch effect on estimates of physical activity, against a criterion measure. In the study by McClain et al (2008) data was collected from 32 school-aged children, (mean (SD) age: 10.3 (0.5) y) over a 30-minute physical education (PE) class and the direct observation tool, the Computerized System for Observing Fitness Instruction Time (C-SOFIT), was used as a criterion measure (McKenzie 2002). The authors reported that shorter epochs resulted in lower individual error in MVPA estimates (McClain et al. 2008).

A limitation of the McClain et al. (2008) study is that the C-SOFIT scale is restricted as a criterion measure because it is only able to provide an aggregate value for the time spent at each intensity level (McClain et al. 2008). The authors recommend that future studies use other direct observation methods which allow for storage of time-series data at the frequency of the minimum epoch under consideration. In addition, the McClain et al. (2008) study looked at a structured PE class which does not necessarily reflect the intermittent free-play activity of younger children. The current study therefore proposed to compare accelerometry estimates of physical activity intensity with the CARS (Puhl et al. 1990) direct observation scale which has been validated for use with young children (Puhl et al. 1990) and calibrated for use at 15-s epochs (Sirard et al. 2005).

The aim of this study was:

- To determine which epoch is most accurate for measuring physical activity in pre-school children during free-play

4.2 METHOD

Video and accelerometry data using the GT1M accelerometer were collected from a sample of 31 pre-school children while they engaged in 1 hour of free-play during their outdoor play-time in their nursery (15 males, 16 females, mean (SD) age: 4.4 (0.8)y, height: 104.8 (6.3) cm, weight: 17.7 (2.5) kg, BMI: 16.1 (1.1) kg/m^2). The mean BMI *z*-score for the sample was 0.20, with 90% classified as 'healthy' weight, 10% classified overweight/obese i.e. BMI at or above 85th centile relative to UK population reference data (Cole 2002). The details of the methods adopted in this study are outlined in the general methods chapter (Chapter 2). The GT1M accelerometers were set to collect data in 1-s epochs.

Data Analysis

Following data collection, the accelerometry data were transferred to an Excel spreadsheet and the 1-s epochs reintegrated into 5-, 15-, 30-, and 60-s epochs. Using a programme in Visual Basic for Applications the data were processed using predetermined age-specific cut-points for physical activity intensity level, as validated with pre-school children by Sirard et al. (2005) for the Actigraph model 7164 (Table 4.1). As discussed in Chapter 3, other cut-points are available, but at the time of data analysis (which was undertaken in 2008 - 9) the Sirard et al. (2005) age-specific cut-points for pre-school children were considered to be most appropriate. To calculate the cut-points for the shorter epochs, cut-points for 60 s were divided by 60, then multiplied to calculate the cut-points for 5-, 15-, 30- and 60-s epochs, as undertaken in earlier studies (Nilsson et al. 2002; Reilly et al. 2008). The number of minutes spent in sedentary behaviour, LPA, VPA and MVPA were calculated. In addition, time spent in total physical activity (TPA) was calculated by combining time spent in light intensity physical activity (LPA) and MVPA.

Table 4.1: Accelerometry cut-points for physical activity and sedentary behaviour.

			Counts/15 s		
Age	**Sed**	**LPA**	**MVPA**	**VPA**	**TPA**
3 year-olds	< 301	≥ 302 to ≤ 614	≥ 615	≥ 1231	≥ 302
4 year-olds	< 363	≥ 364 to ≤ 811	≥ 812	≥ 1235	≥ 364
5 year-olds	< 398	≥ 399 to ≤ 890	≥ 891	≥ 1254	≥ 399

LPA: light physical activity; MVPA: moderate-to-vigorous activity; Sed: sedentary behaviour, TPA: total physical activity; VPA: vigorous physical activity.

(Sirard et al. 2005)

Normality tests were conducted using the Shapiro-Wilks statistic because the sample was less than 50 and the data were found to be non-normally distributed ($p < 0.001$) (Appendix IV, Appendix Table IV.i). Descriptive data on the median values and IQR values were calculated and are presented.

The CARS scale (Puhl et al. 1990) was used in this study as the criterion measure to directly observe and score the children's physical activity during free-play. The rationale and the method adopted for scoring the CARS is outlined in the general methods chapter (Chapter 2). The CARS was coded in 15-s epochs and this allowed comparison to be made with accelerometry data from 15-, 30- and 60-s epochs. The Wilcoxon signed rank test, which is the non-parametric equivalent of the paired *t*-test, was used to compare time spent in sedentary behaviour and different intensities of physical activity against the CARS at 15-s and 60-s epochs. A Bonferroni correction was applied to adjust the level of significance to reduce the risk of a Type I error (α/number of comparisons). Bland and Altman plots (Bland and Altman 1986) were used to compare accelerometry estimates of physical activity against the CARS for data collected at 15- and 60-s epochs.

4.3 RESULTS

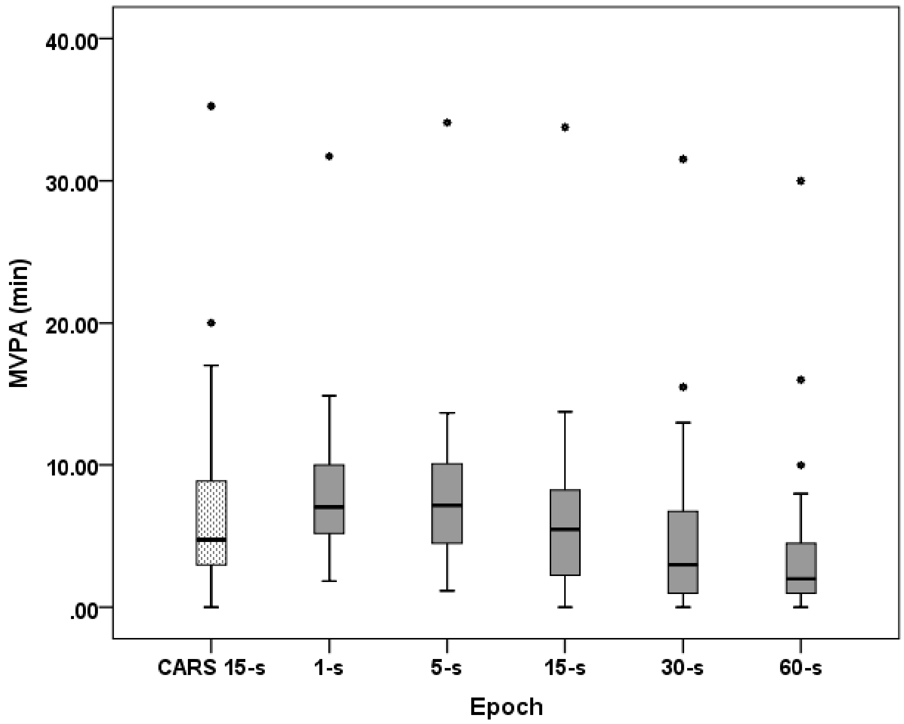
Table 4.2 presents the median (IQR) of the minutes spent at LPA, MVPA and of the time spent in sedentary behaviour when different epoch lengths are used. Figure 4.1 illustrates the differences in the median (IQR) minutes of time spent in MVPA when different epoch lengths are used.

Table 4.2: Median (IQR) minutes of time spent in different intensities (Sirard et al, (2005) cut-points).

	Epoch (seconds)				
	1 s	**5 s**	**15 s**	**30 s**	**60 s**
Sed	33.6 (14.5)	30.2 (13.7)	29.8 (14.5)	27.0 (16.0)	25.0 (19.0)
LPA	5.7 (3.8)	8.3 (4.7)	12.5 (7.3)	14.5 (10.0)	15.0 (15.0)
VPA	2.9 (2.7)	1.5 (2.3)	0.5 (1.3)	0.0 (1.0)	0.0 (0)
MVPA	7.1 (4.9)	7.2 (5.6)	5.5 (6.0)	3.0 (6.0)	2.0 (4.0)
TPA	13.9 (7.0)	15.8 (9.4)	16.7 (10.8)	19.0 (13.5)	18.0 (15.0)

LPA: light physical activity; MVPA: moderate-to-vigorous activity; Sed: sedentary behaviour, TPA: total physical activity; VPA: vigorous physical activity.

Figure 4.1: Box plot of median (IQR) minutes for time spent in MVPA using different epochs.



In Figure 4.1 can be seen that with shorter epochs more minutes are classified as MVPA.

Table 4.3 presents the median (IQR) minutes for time spent in sedentary behaviour, LPA, VPA, MVPA and TPA when 15- and 60-s epochs are used, compared against the CARS criterion measure (Puhl et al. 1990).

Table 4.3: Median (IQR) minutes spent at different intensities comparing 15- and 60-s epochs.

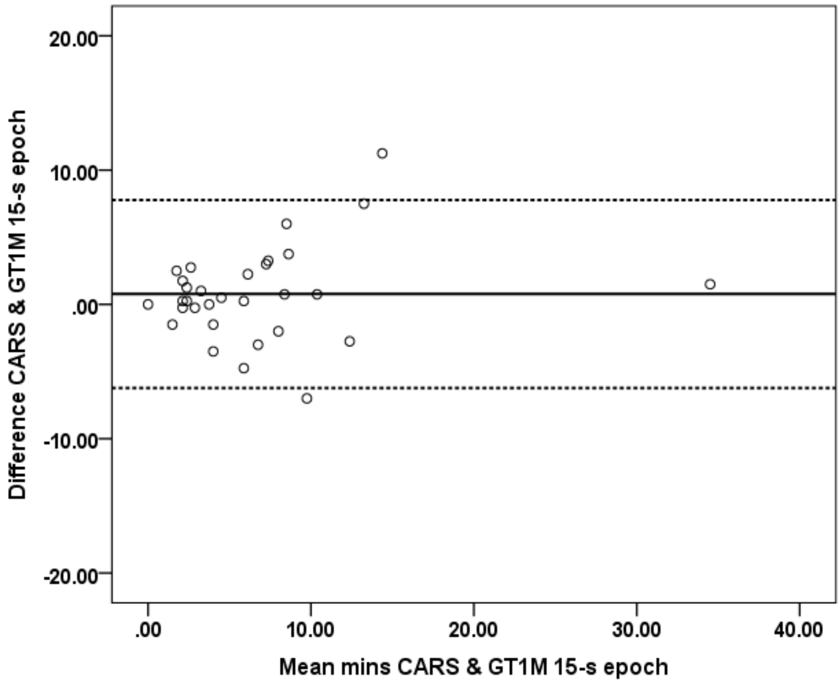
	Epoch (seconds)		
Intensity level	**CARS 15 s**	**15 s**	**60 s**
Sed	12.0 (11.3)	29.5(14.5)*	25.0 (19.0)*
LPA	27.8 (14.8)	12.5 (7.3)*	15.0 (15.0)*
VPA	0.3 (1.8)	0.5 (1.3)	0.0 (0.0)
MVPA	4.8 (6.0)	5.5 (6.0)	2.0 (4.0)*
TPA	33.0 (13.5)	16.8 (10.8)*	18.0 (15.0)*

**significant difference between estimate and CARS value (p < 0.001); IQR: inter quartile range; LPA: light physical activity; MVPA: moderate-to-vigorous activity; Sed: sedentary behaviour; TPA: total physical activity; VPA: vigorous physical activity.*

There were significant differences between the number of minutes classified by the CARS as sedentary behaviour, LPA and TPA and the accelerometry estimates ($p < 0.001$). It is interesting to note that more time was classified as sedentary and less time as LPA with the accelerometry estimates in comparison to the CARS. The median (IQR) number of minutes classified as MVPA by CARS was 4.8 (6.0) min, while the median MVPA by the GT1M ranged from 2.0 (4.0) min at 60-s epoch to 5.5 (6.0) min at 15-s epochs. There was no significant difference between accelerometry estimates of MVPA collected in 15 s compared with the CARS ($z = -1.27$, $p = 0.21$, $r = 0.23$).

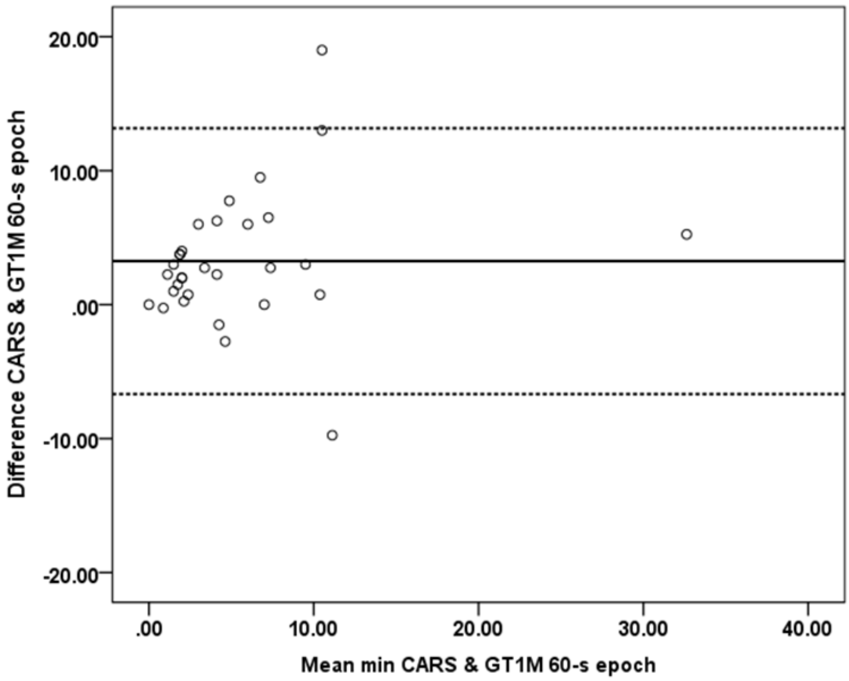
Bland and Altman plots were undertaken to explore agreement between estimated minutes of MVPA using 15- and 60-s epochs, compared against the CARS criterion measure (Figure 4.2 and Figure 4.3 respectively). The mean difference (LOA) between the number of minutes recorded by the GT1M and direct observation (CARS Level 4 and 5) was 0.8 (-6.2 to 7.8) min at 15-s epochs and 3.2 (-6.8 to 13.2) min at 60-s epochs. It can been seen that when using 60-s epochs there were wider limits of agreement than with the 15-s epochs and the difference appears to increase in magnitude as estimated time spent in MVPA increases. Both plots reveal heteroscedasticity, suggesting data were randomly distributed, with no systematic bias (Corder et al. 2007).

Figure 4.2: Bland and Altman plot of accelerometry estimates of MVPA at 15-s epoch against the CARS.



MVPA: moderate-to-vigorous activity. Solid line depicts mean difference and the dotted line the limits of agreement (LOA). Mean difference (LOA) 0.8 (-6.2 to 7.8) min.

Figure 4.3: Bland and Altman plot of accelerometry estimates of MVPA at 60-s epoch against the CARS.



MVPA: moderate-to-vigorous activity. Solid line depicts mean difference and the dotted line the limits of agreement (LOA) Mean difference (LOA) 3.25(-6.8 to 13.2) min.

4.4 DISCUSSION

The aim of this study was to determine which epoch is most accurate for measuring physical activity in pre-school children during free-play. Similar to earlier studies, the results from this study suggest there is an epoch effect, with shorter epochs recording greater numbers of minutes of MVPA (Edwardson and Gorely 2010; Mahar et al. 2008; Reilly et al. 2008; Vale et al. 2009). There was a 5.1 minute difference in estimates of the median time spent in MVPA when 1-s epochs were compared with 60-s epochs. This may be as a result of the smoothing effect whereby the shorter bouts of MVPA activity are averaged with longer bouts of lower intensity activity over an epoch, resulting in activity for the epoch being misclassified as low intensity (McClain et al. 2008).

In the current study there was, however, good absolute agreement between the 15-s epoch estimates of time spent in MVPA and the CARS criterion measure at 15 s. It was not possible from the methods adopted in this study to determine if shorter epochs, e.g. less than 15 s, would offer greater accuracy for assessing intensity of physical activity, as the CARS was scored in 15-s periods. Later studies have scored the CARS using second-by-second coding to allow comparison with accelerometry data collected at shorter epochs (Oliver et al. 2009). However this approach to using the CARS has not been validated and it is possible that coding observations in 1-second intervals could increase the risk of human error. While a difference of 5.1 minutes of time spent in MVPA between 1- and 60-s epoch may not seem a large amount of time, if extrapolated over a day this could result in large differences in estimates of MVPA and it is still possible that 15-s epochs could result in an underestimation of time spent in MVPA. Recent studies have recommended that shorter epochs of 1 to 5 s should be used in studies of young children, although there is no empirical evidence to support these shorter epochs (Ojiambo et al. 2011). It is, however, argued that shorter epochs offer greater sensitivity to detect changes in the intensity of physical activity, as well as to detect the brief interruptions (rest during high intensity activity) (Ayabe et al. 2013).

While there was a non-significant difference in time spent in vigorous physical activity using the different epochs, there were very few minutes of vigorous activity recorded over the 1-hour period (median (IQR): 0.5 (1.3) min with 15-s epoch). While this vigorous time may be misclassified as moderate intensity physical activity, it seems more likely that moderate intensity activity was being misclassified as LPA (median (IQR): 1.0 (3.0) min at 60-s epoch compared to 4.5 (4.5) min at 15-s epoch). The results of this study suggest that for sedentary behaviour, LPA and TPA there was a significant difference between the accelerometry estimates and the CARS at all epoch lengths. This may be explained in part by the cut-points adopted to classify the data into different intensities. The concern with the differences in estimates of physical activity intensity resulting from the application of different accelerometry cut-points is highlighted in the study discussed in Chapter 6 of this thesis.

In addition to the concern over which cut-points are accurate, is that the Sirard et al. (2005) cut-points have been developed and validated for the 7164 accelerometer, which has a different internal technology from the GT1M accelerometer (John et al. 2010). It is unclear whether these cut-points are appropriate or valid for the GT1M and this methodological concern will be the focus of Chapter 6. Finally, it is unclear whether these differences may in part be due to the sporadic nature of physical activity seen in younger children and whether triaxial accelerometers offer a more accurate means of measuring their physical activity. This will be the focus of the next chapter.

In conclusion, there is good absolute agreement between accelerometry estimates of MVPA at 15-s epochs and the criterion measure of direct observation. The results suggest that grouping the time spent in moderate and vigorous intensity physical activity together into one category of MVPA possibly addresses the concern with vigorous intensity activity being misclassified as moderate intensity (Reilly et al. 2008). However in this study, time spent in moderate intensity physical activity is also possibly being misclassified as light physical activity when longer epochs are used.

CHAPTER 5 : COMPARISON OF TRIAXIAL VERSUS UNIAXIAL ACCELEROMETRY DURING FREE-PLAY IN PRE-SCHOOL CHILDREN

5.1 INTRODUCTION

This chapter is based on the researcher's study, Hislop et al. (2012b), published in Pediatric Exercise Science in August 2012.

Accelerometers are sensors which measure acceleration of an object, such as the body, along one or more reference axis (Yang and Hsu 2010). Different types of accelerometers are available, including uniaxial accelerometers, which measure acceleration in the vertical plane, omnidirectional accelerometers, which are sensitive to motion in any direction (Chen and Bassett 2005), and triaxial accelerometers, which have the ability to measure acceleration in three orthogonal planes (vertical, anteroposterior, mediolateral) and can provide a vector magnitude value for acceleration (Vanhees et al. 2005).

The majority of accelerometry studies with pre-school children have used the Actigraph uniaxial accelerometers which are positioned so as to measure acceleration of the body in the vertical plane (Cain et al. 2013). The accuracy of these studies has been called into question because unlike triaxial accelerometers, uniaxial accelerometers are unable to detect movements in all three planes (Trost et al. 2005). It has been argued that since pre-schoolers participate in activities that require less vertical movement and more omnidirectional movement, triaxial accelerometers might be better able to capture and characterise their movement (Tanaka and Tanaka 2009). It would therefore theoretically be preferable to use triaxial accelerometry rather than uniaxial accelerometry. However, few studies with children have carried out formal comparisons of triaxial versus uniaxial devices (Eston et al. 1998; Ott et al. 2000; Welk et al. 2000a) and there is no conclusive evidence that one model is superior to another (Rowlands 2007).

In one of the first studies to compare triaxial with uniaxial accelerometers in children, Eston et al. (1998) found that the Tritrac-RD3 triaxial accelerometer (Professional products, Reining, Madison, WI, USA) was more accurate at predicting oxygen uptake than the 7164 uniaxial accelerometer in 30 children (mean (SD) age: 9.2 (0.8) y) who undertook treadmill walking and running, as well as structured activities. However, a possible limitation of this study is that it was laboratory based and did not examine unstructured free-living physical activity characteristics of young children. In the Eston et al. (1998) study both the uniaxial and the Tritrac-R3D accelerometers collected data in 1-min epochs which, as discussed in Chapter 3 and Chapter 4, may not accurately capture the physical activity behaviour of pre-school children. Finally, the Tritrac-R3D was a rather bulky monitor (120 x 65 x 22 mm, 168 g), which needed to be taped in place to prevent unwanted movements and it is not clear whether these factors mean that younger children may not have tolerated wearing these accelerometers over an extended period of time. The Tritrac-R3D has, however, subsequently been replaced by the much smaller, more user-friendly RT3 accelerometer (71 x 56 x 28 mm, 65.2g), which can collect data in 1-s epochs (Powell and Rowlands 2004). While Rowlands et al. (2004) have validated the RT3 against a criterion measure of oxygen uptake in school-aged children (n = 19, mean (SD) age: 9.5 (0.8) y), they reported that the output from the RT3 was not comparable with the Tritrac-R3D. They found that the threshold count for the RT3 was significantly higher ($p < 0.05$) than for the Tritrac-R3D for a variety of activities, including treadmill walking and running, as well as kicking a ball and playing hopscotch.

To date, the only study with pre-school children which has set out to compare a uniaxial accelerometer with an accelerometer which can measure in different planes is by Kelly et al. (2004). In this study, the 7164 uniaxial accelerometer was compared with the Actiwatch, an omnidirectional accelerometer, both of which were worn simultaneously by 78 three- to four-year-old children who were observed in free-play for 45 minutes. Accelerometry outputs were compared against the CPAF direct observation measure (O'Hara et al. 1989). The results suggested that the uniaxial accelerometer had greater accuracy as the output was significantly positively correlated with the CPAF for total physical activity ($r = 0.72$, $p < 0.05$). In contrast, the Actiwatch was not correlated with the CPAF ($r = 0.16$, $p > 0.05$) (Kelly et al. 2004). However, it is argued that while the sensor in the Actiwatch is omnidirectional, it is most sensitive to acceleration in the vertical plane (Chen and Bassett 2005) and functions as a single axis device (Pate et al. 2010). Therefore, whether it is advantageous to use accelerometers which measure acceleration in different planes simultaneously for the measurement of physical activity in pre-school children has still to be determined.

Triaxial accelerometers, particularly the RT3 monitor, have been widely used in studies with school-aged children (Chu et al. 2007; Eston et al. 1998; Hoos et al. 2004; Hussey et al. 2009; Hussey et al. 2009; Louie et al. 1999; Ott et al. 2000; Rowlands et al. 1998; Rowlands et al. 2004; Rowlands and Eston 2005; Sun et al. 2008) and adolescents (Vanhelst et al. 2010a; Vanhelst et al. 2010; Vanhelst et al. 2012). However, only three triaxial accelerometry studies have been conducted with pre-school children (Tanaka et al. 2007; Tanaka and Tanaka 2009; Tanaka et al. 2012) and these used the ActivTracer (GMS, Tokyo, Japan) triaxial accelerometer. In the first study, linear and non-linear equations were used to estimate EE and physical activity ratio (PAR) from the ActivTracer data using indirect calorimetry as the criterion measure. Data were collected from 27 pre-school children (mean (SD) age: 0.6 (0.3) y) who engaged in nine different activities. The accelerometry count (expressed as m*G*) was calculated as the change in absolute values for acceleration in each direction (vertical, anteroposterior, mediolateral). In addition 'synthetic' (synthesised tri axes vector) acceleration counts were obtained during the nine movements. The ActivTracer accelerometer was found to provide valid measures of EE and PAR. The threshold between light and moderate activity was identified as being 395 m*G*. The sensitivity and specificity of this threshold were 77% and 94% respectively to distinguish between light and moderate intensity physical activity with the 'synthetic' and vertical/horizontal acceleration counts. In the authors' subsequent study the ActivTracer was used to provide concurrent validity for pedometry step counts for MVPA with data collected from 212 pre-school children (4 to 6 y) over 6 days. In this second study the authors used 130 to 600 m*G* as the accelerometry threshold for MVPA. While these are the only studies to date to have published cut-points for triaxial accelerometry for pre-school children, unfortunately the Tanaka et al. (2007) cut-points are not applicable for use with the RT3 accelerometer.

Despite the availability of the RT3 and its wide use in studies of school-aged children and adolescents, its validity has neither been investigated with pre-school populations, nor has a comparison of the RT3 with uniaxial accelerometry been undertaken to determine whether triaxial offers a more accurate means of quantifying physical activity behaviour. It is therefore unclear whether there are advantages from using a triaxial accelerometer to measure physical activity of pre-school children. Finally, given the growing number of studies which are using triaxial accelerometry, it would also be beneficial to know whether the uniaxial and triaxial devices give similar information, to allow comparison between outcomes of different studies.

The aim of this study was:

- To investigate whether there are advantages to using triaxial over uniaxial accelerometry to measure physical activity in pre-school children during free-play.

5.2 METHOD

Video and accelerometry data were collected from a convenience sample of 31 pre-school children while they engaged in 1 hour of free-play during unstructured play-time in their nursery (15 males, 16 females, mean (SD) age: 4.4 (0.8) y, height: 104.8 (6.3) cm, weight: 17.7 (2.5) kg, BMI: 16.1 (1.1) kg/m^2). The sample had a mean BMI z-score of 0.20 and 90% were classified as 'healthy' weight and 10% classified as overweight/obese i.e. BMI at or above 85th centile relative to UK population reference data (Cole 2002). In this study measurements were taken simultaneously from two accelerometers: the GT1M Actigraph uniaxial model and the RT3 triaxial accelerometer. The RT3 accelerometer measures acceleration in the three orthogonal planes: vertical, anteroposterior, mediolateral. The RT3 was set to provide vector magnitude data which combines data from all three axis of motion. The raw data from both accelerometers were filtered and digitised, converting it to activity counts over a predefined period (epoch). These activity counts can be compared against predetermined cut-points for intensity levels. In the current study the epoch was set at 1 s. The details of the method for this study are outlined in the general methods chapter (Chapter 2).

Data from the accelerometers were transferred to an Excel spreadsheet and the 1-s epoch reintegrated into 15-s epochs (Edwardson and Gorely 2010). Using a Visual Basic for Applications programme in Excel, the data were processed using age-specific cut-points for MVPA intensity level, as validated with pre-school children by Sirard et al. (2005) for the 7164 accelerometer and applied to the GT1M data in this study (GT1M^s). Several other MVPA cut-points have been published for the Actigraph for young children, ranging from 368 cpm to 3200 cpm (Table 5.1) and the data were also processed using these cut-points to allow comparison.

While there are several published cut-points for the RT3 for older children (Table 5.1), many have been developed from calibration studies which have used three METs as the threshold for MVPA. It is recognised that applying adult MET values may not be appropriate, as children's resting metabolic rates are higher (Ridley and Olds 2008). Sun et al. (2008) validated cut-points for several moderate free-living activities, including kicking and catching a ball, walking and jogging. Two of the cut-points suggested by Sun et al. (2008) were selected for the current study: walking relaxed (RT3^{WR}; counts for MVPA > 413 counts/15 s) and light jog (RT3^{LJ}; counts for MVPA > 780 counts/15 s). Data processing was undertaken to allow cross comparison between minutes of MVPA resulting from the application of the different RT3 cut-points.

Table 5.1: Summary of calibration studies of cut-points for the RT3, and Actigraph (GT1M & 7164 models) accelerometers in children.

Accelerometer model	Authors	No. participants (n), age range, mean (SD)	Criterion measure	Activities	Criterion for MVPA intensity	MVPA, counts/15 s
RT3	Chu et al. (2007)	35 8 - 12 y, 11.1 (1.0) y	VO_2 (portable metabolic unit)	TM W, R	≥ 3 METs	≥ 465
RT3	Rowlands et al. (2004)	19 9.5 (0.8) y; Boys only	VO_2 (portable metabolic unit)	TM W, R, other	≥ 3 METs	≥ 243
RT3	Sun et al. (2008)	25 indoor; 8 outdoor 12 - 14 y	VO_2 (portable metabolic unit)	Sit, cycle, TM W, R, FL	WR LJ	WR ≥ 413 LJ ≥ 780
RT3	Vanhelst et al. (2010b)	40 10 - 16 y	VO_2 (portable metabolic unit)	TM W, R	TM W, R at ≥ 3 $km \cdot h^{-1}$	≥ 238
7164	Evenson et al. (2008)	33 5 - 9 y, 7.3 (1.1) y	VO_2 (portable metabolic unit)	W, R, other	TM brisk W, R ≥ 3 mph, dribble basketball, stair climbing, jumping jacks.	≥ 574
7164	Freedson et al. (1997)	80 6 - 18 y	VO_2 (portable metabolic unit)	TM W, R	≥ 3 METs	3 y ≥ 92 4 y ≥ 111 5 y ≥ 133
7164	Pate et al. (2006)	29 3 - 5 y, 4.4 (0.8) y	VO_2 (portable metabolic unit)	Rest, TM W, R, FL	TM brisk walking	≥ 420
7164	Puyau et al. (2002)	26 6 - 16 y, 10.7 (2.9) y	VO_2 (room calorimetry)	W, R, FL	TM walk, run at ≥ 3.5 mph (6 - 7 y olds); 4.5 mph (8 - 16 y olds). MVPA free-play	≥ 800
7164	Sirard et al. (2005)	16 3-, 4-, 5- y	CARS	Sit, play, W, R	Fast W, run ≥ 4.3 ± 0.6 $km \cdot h^{-1}$	3 y ≥ 615 4 y ≥ 812 5 y ≥ 891
GT1M	Van Cauwenberghe et al. (2011)	18 4 - 6 y, 5.8 (0.3) y	CARS	TM W, R FL W	Brisk W ≥ 4.8 $km \cdot h^{-1}$. CARS score 3.1 - 4.0	≥ 585

CARS: Children's Activity Rating Scale; FL: free-living; LJ: light jog; MVPA: moderate-to-vigorous activity; R: run; TM: treadmill; W: walk; WR: walking relaxed.

Analysis

Data were imported into SPSS (version 17) for analysis. Normality tests were conducted using the Shapiro-Wilks statistic since the sample was less than 50. The count per minute (cpm) data were found to be normally distributed ($p > 0.05$); and the mean and SD were reported. All other data on minutes spent in MVPA were found to be not normally distributed (Appendix V, Appendix Table V.i). As a result the median and interquartile range values (IQR) were reported for these variables. Pearson's correlation coefficient was calculated to explore the relationship between total cpm for each accelerometer. Spearman's rank correlation was calculated to determine whether the different approaches provided a similar relative assessment of MVPA when compared against the CARS score.

Using the Friedman's Repeated Measures ANOVA the difference between the number of minutes of MVPA recorded by each accelerometer model and direct observation at 15-s epochs was explored, with post hoc analysis using the Wilcoxon signed-rank test. To reduce type I error a Bonferroni correction was applied so that the significance level was set at $p < 0.01$. Finally, Bland and Altman plots with limits of agreement were plotted to explore the relationship between the accelerometry estimates of MVPA and the CARS criterion measure of MVPA (Bland and Altman 1986).

5.3 RESULTS

During the 1 hour of free-play the mean (SD) of total counts per minute (cpm) for the RT3 was 1544 (442) cpm and for the GT1M was 1300 (476) cpm. There was a significant positive correlation ($r = 0.72$, $p < 0.001$) in the total counts between the RT3 and the GT1M accelerometers. Spearman's rank order correlations between accelerometry cpm and percentage time spent in MVPA as recorded by direct observation for the GT1M was $r = 0.56$ ($p < 0.01$) and for the RT3 $r = 0.39$ ($p < 0.03$).

Table 5.2 presents the median (IQR) of minutes of MVPA at 15-s epochs resulting from the application of the different published cut-points for the RT3 and GT1M accelerometers.

Table 5.2: Comparison of median (IQR) minutes of MVPA as classified by cut-points.

Authors	Accelerometer model/CARS	Median (IQR) min MVPA
Puhl et al.	CARS	4.8 (6.0)
Vanhelst et al.	RT3	27.5 (12.8)
Rowlands et al.	RT3	27.3 (12.3)
Sun et al.	$RT3^{WR}$	17.0 (10.3)
Sun et al.	$RT3^{LJ}$	5.8 (5.8)
Chu et al.	RT3	15.3 (9.0)
Freedson et al.	GT1M	31.0 (12.3)
Pate et al.	GT1M	14.3 (11.8)
Evenson et al.	GT1M	9.0 (9.5)
Van Cauwenberghe et al.	GT1M	8.8 (9.3)
Sirard et al.	$GT1M^{s}$	5.5 (6.0)
Puyau et al.	GT1M	4.3(6.0)

CARS: Children's Activity Rating Scale; $GT1M^{s}$: Sirard et al. (2005) cut-points; $RT3^{WR}$: walking relaxed; $RT3^{LJ}$: light jog.

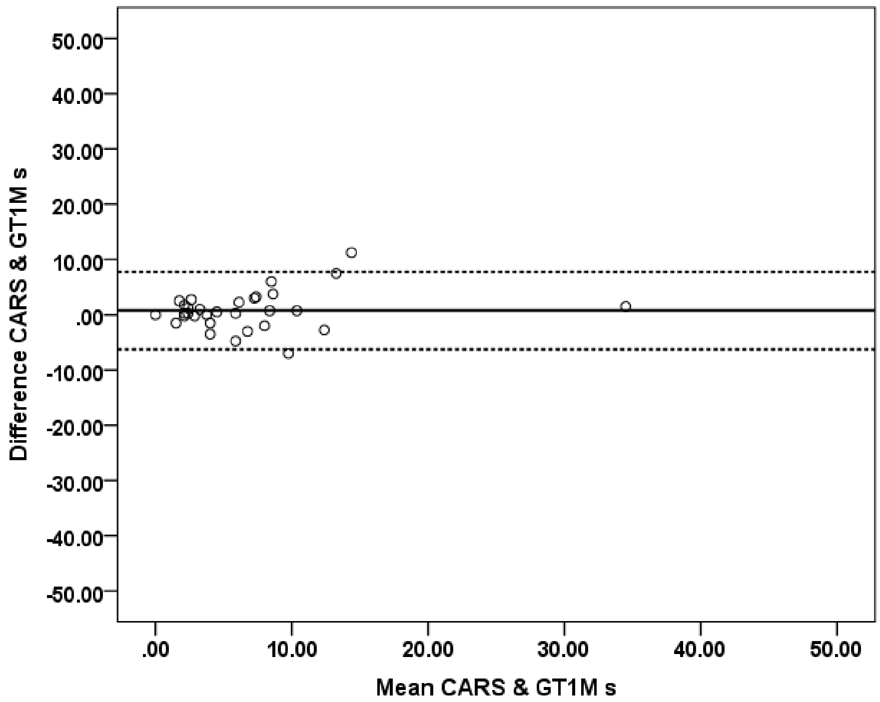
Using the Friedman's Repeated Measures ANOVA there was a significant difference between the number of minutes of MVPA at 15-s epoch between the $GT1M^{s}$, $RT3^{LJ}$, $RT3^{WR}$ and the CARS score (χ^2 (3) = 58.5, $p < 0.05$). Post hoc analysis with the Wilcoxon test using a Bonferroni correction revealed that there was a non-significant difference between the number of minutes classified as MVPA using the $GT1M^{s}$ (*Mdn* = 5.5) and the CARS (*Mdn* = 4.8, z = -1.266, p = 0.206) and the CARS (*Mdn* = 4.8) and the $RT3^{LJ}$ (*Mdn* = 5.8, z = -7.46, p = 0.456) and between the $GT1M^{s}$ (*Mdn* = 5.5) and the $RT3^{LJ}$ (*Mdn* = 5.8, z = -0.844, p = 0.399). There was a significant difference between the CARS (*Mdn* = 4.8) and the $RT3^{WR}$ (*Mdn* = 17.0, z = -4.861, p < 0.00).

Bland and Altman plots were undertaken to explore agreement between the mean number of minutes of MVPA determined using the CARS criterion measure at 15-s epochs and the $GT1M^{s}$ data, the $RT3^{LJ}$ and the $RT3^{WR}$ (Figure 5.1). The mean difference (LOA) between the number of minutes recorded by GT1M and direct observation (CARS Level 4 and 5) was 0.8 (-7.7 to 6.3) min. The mean difference (LOA) for the $RT3^{WR}$ was -12.3 (-27.4 to 3.2) min and for the $RT3^{LJ}$ was 0 (-9.0 to 9.0) min.

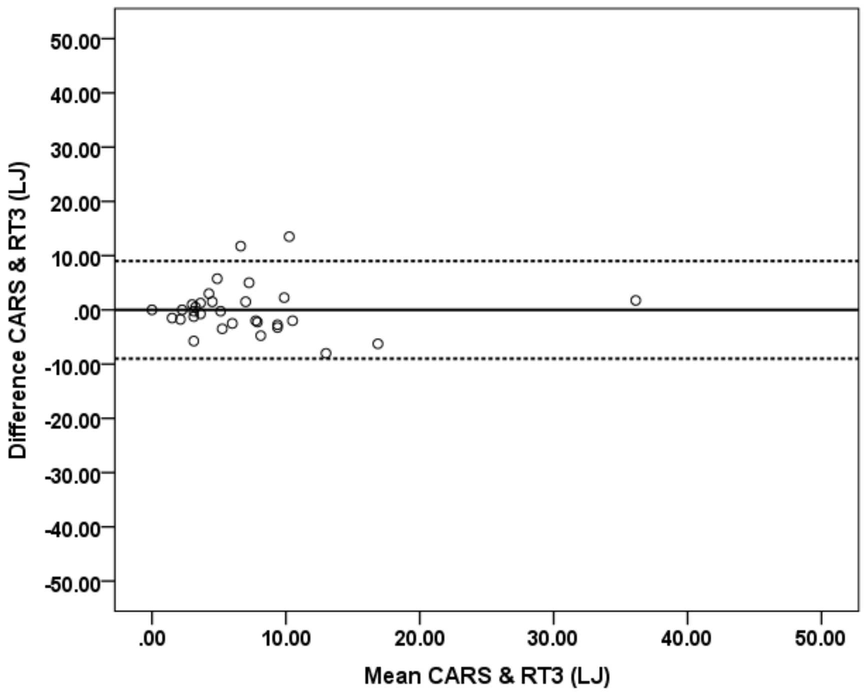
.

Figure 5.1: Bland and Altman Plots of MVPA (i) between CARS and GT1M^{s} (ii) between CARS and RT3^{LJ} (iii) between CARS and RT3^{WJ} at 15-s epoch.

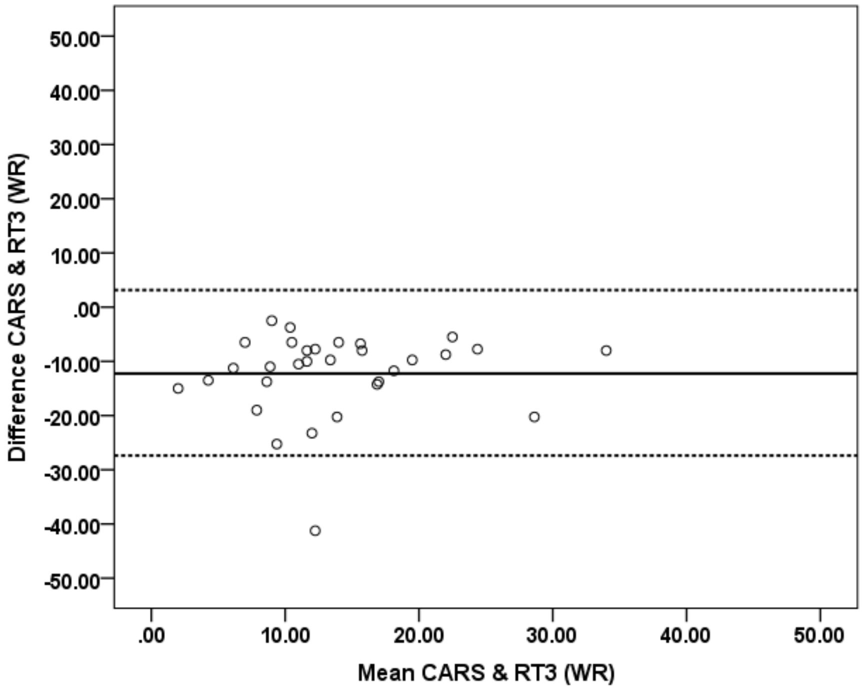
(i)



(ii)



(iii)



CARS: Children's Activity Rating Scale; WR: walking relaxed; LJ: Light Jog. The solid lines depict the mean difference in time and the dotted lines indicate the limits of agreement (LOA).

5.4 DISCUSSION

The aim of the current study was to compare the GT1M uniaxial accelerometer with the RT3 triaxial accelerometer to determine whether the RT3 triaxial accelerometer is more accurate in capturing physical activity levels of pre-school children. The study found no evidence that the RT3 triaxial accelerometer is more accurate than the GT1M uniaxial accelerometer for estimates of absolute amount of time spent in MVPA.

A significant positive correlation in the accelerometry output between the RT3 and GT1M models (r = 0.72) was found and the Spearman's rank order correlation with percentage time spent in MVPA as measured by the CARS was significant. However, the correlations were not strong for either accelerometer (GT1M r = 0.56 and RT3 r = 0.39), suggesting possible limitations in the relative assessment of MVPA if raw accelerometry count output is used.

The results of this study illustrate the problems that occur as a consequence of applying different cut-points to accelerometry data. The cut-points for Actigraph accelerometers range from 368 cpm (Freedson et al. 1998), which would result in median (IQR) of 31(12.3) min of MVPA, to 3200 cpm (Puyau et al. 2002) cut-points, which would result in 4.3 (6.0) min of MVPA during 1 hour of free-play. The existence of different cut-points has been attributed to their being developed from studies which have used different calibration methods, such as treadmill and free-living protocols, different criterion measures, such as indirect calorimetry and direct observation, and the use of three METs as the threshold for MVPA in children which may be too low (Guinhouya et al. 2006). A limitation of the current study is that many of the available cut-points for the RT3 are based on a threshold of three METs and range from 952 cpm (Vanhelst et al. 2010b), resulting in 27.5 (12.8) min of MVPA, to 1860 cpm (Chu et al. 2007), resulting in 16.8 (9.8) min of MVPA. As Sun et al. (2008) do not provide a definitive cut-point for MVPA, but list moderate intensity activities and their respective cut-points, a decision was made to select two of these moderate intensity activities. Other 'moderate' cut-points were not evaluated and further studies are therefore required to validate appropriate cut-points for the RT3 for pre-school children.

The $RT3^{WR}$ cut-point appeared to overestimate the number of minutes of MVPA in pre-school children (median (IQR): 19.3 (10.3) min) when compared against direct observation (median (IQR): 4.8, (6.0) min). The $RT3^{LJ}$, however, was more accurate (median (IQR): 7.1 (6.9) min). This may simply be a reflection of the discrepancy between validation studies as to what activities constitute moderate intensity activity. In the original CARS study by Puhl et al (1990), slow walking was categorised as a light intensity activity and was coded as a Level 3 activity. In the study by Van Cauwenberghe et al. (2011), moderate intensity activity was classified as a CARS score of greater than 3.1. Using this approach would result in classifying slow walking as moderate intensity activity and in the original development of CARS by Puhl et al. (1990) slow walking corresponded to an energy expenditure of less than three times individual children's resting metabolic rate. In the current study only those activities coded as Level 4 and greater were classified as moderate intensity activity, for example, fast walking. This may partially explain why differences in proposed cut-point thresholds exist.

The current study made use of the GT1M accelerometer, which is a more recent Actigraph model than the 7164 model, from which the Sirard et al. (2005) cut-points were developed and validated. While some papers have found good cross-validation between this model in a laboratory setting (John et al. 2010; Kozey et al. 2010b), other studies state that the GT1M may be underestimating levels of physical activity suggesting the application of a correction factor (Corder et al. 2007). Although not presented in the results of this study, it is noted from post hoc analysis that the application of the + 9% correction factor recommended by Corder et al. (2007) to the GT1M^s data would result in positive bias (LOA) of 0.2 (-7 to 7.4) min, slightly less than the 0.8 min presented for the uncorrected GT1M^s data. This could suggest that the correction to the GT1M output recommended by Corder et al. (2007) might improve the accuracy of the GT1M measurement of MVPA, at least in young children. This comparability of output between different generations of Actigraph uniaxial accelerometers will be explored further in Chapter 7.

The present study suggests that the GT1M may have good absolute validity when compared against the CARS criterion measure, with a small bias, indicating accuracy particularly for group assessments of MVPA. However, for assessments of individual levels of MVPA, the current study is less supportive of absolute accuracy, given the large limits of agreement (-7.7 to 6.3 min) between the CARS and the GT1M^s estimates. The average of approximately 6 minutes of MVPA recorded over a 1 hour of free-play session in the current study might seem relatively low. However, if sustained over the whole day this would lead to accumulation of more than 1 hour of MVPA per day, and in fact the levels of MVPA in the current study are higher than those observed in most previous nursery-based studies (Reilly 2010). It should be noted that the usefulness of MVPA has been questioned as a concept for the pre-school age-group, for example, the recent UK guidance on physical activity for early years emphasises total volume of physical activity and does not recommend an amount of time to be spent in MVPA (Department of Health, Physical Activity, Health Improvement and Protection 2011).

In conclusion, the current study suggests that there is no advantage to using triaxial accelerometry, at least with the RT3, over a uniaxial accelerometer, for studies of pre-school children, in the assessment of either relative or absolute amounts of physical activity. The RT3 cut-points by Sun et al. (2008) for light jog provide a threshold for MVPA in this age group which has good absolute agreement for MVPA against the CARS criterion measure. Comparability of output from studies which have used triaxial or uniaxial accelerometers may be problematic due to widely different thresholds for categories of physical activity intensity being adopted. The concern about which cut-point is accurate for the GT1M is the focus of the next chapter.

CHAPTER 6 : COMPARISON OF ACTIGRAPH ACCELEROMETRY CUT-POINTS FOR PHYSICAL ACTIVITY AND SEDENTARY BEHAVIOUR IN PRE-SCHOOL CHILDREN: A VALIDATION STUDY

6.1 INTRODUCTION

This chapter is based on the researcher's validation study of accelerometry thresholds, Hislop et al. (2012a) published in Pediatric Exercise in 2012.

One of the outstanding methodological questions regarding accelerometry data processing and interpretation is how to transform the accelerometry count output into a biologically meaningful format (Cliff et al. 2009b). One approach is to apply count thresholds, also referred to as cut-points, to the accelerometry count output to classify the data into the different intensity levels: sedentary behaviour, light, moderate and vigorous intensity physical activity. This approach allows for the determination of the number of minutes spent at these different intensities. There is, however, a lack of consensus over which cut-points should be used (Kim et al. 2012). Numerous cut-points for Actigraph accelerometers have been developed for children and some for pre-school children and there are a variety of cut-points available. For example, the cut-points for moderate-to-vigorous activity (MVPA) range from activity above 1263 cpm (Freedson et al. 1997) to activity above 3600 cpm (Mattocks et al. 2007). Different cut-points for time spent in sedentary behaviour have also been calibrated for children and these range from $\leq$ 100 cpm (Evenson et al. 2008) to < 1592 cpm (Sirard et al. 2005). The application of different cut-points makes comparison between studies problematic, leading to conflicting conclusions about levels of sedentary behaviour, MVPA and total physical activity (TPA) and without agreement on cut-points it is difficult to ascertain whether children are meeting physical activity guidelines or not (Beets et al. 2011a; Kim et al. 2012).

As discussed in the introductory chapter, discrepancies in cut-points may in part be due to the differing criterion methods and protocols used during calibration studies, such as direct observation (Sirard et al. 2005) and indirect calorimetry (Pate et al. 2006), assessed while children have engaged in treadmill-based, or free-living activities, or both. Many studies have used linear regression equations to calibrate cut-points against energy expenditure, which may also be problematic as output from Actigraph accelerometers at higher intensities has been found not to increase linearly (Brage et al. 2003). The application of equations developed for adults is not appropriate for children and the use of three METs as a threshold for MVPA activity in young children could result in an overestimation of time spent in MVPA (Guinhouya and Hubert 2008). To further complicate matters, there is a lack of agreement on definitions of sedentary behaviour, LPA and MVPA in pre-school children. Sedentary behaviour is either defined as predominantly sitting (Pate et al. 2010), or to include standing and sitting activities (Martin et al. 2011), while MVPA can include slow walking (Pate et al. 2006). In other studies, slow walking is defined as a light intensity activity (Puhl et al. 1990; Sirard et al. 2005). Finally, some researchers argue that age-specific cut-points are required (Sirard et al. 2005), while other researchers suggest that these are not needed (Evenson et al. 2008); this issue remains unresolved.

A recent study has compared cut-points for youth (Trost et al. 2011) and studies by Cliff and Okely (2007) and Guinhouya et al. (2006) have highlighted discrepancies in quantification of MVPA when different cut-points are applied. However, the crucial question of which cut-point is most accurate compared against an external criterion method, has yet to be answered for pre-school children. Direct observation is recognised as a criterion method for measuring physical activity which is particularly suited to studies of young children (Freedson et al. 2005). The Children's Activity Rating Scale (CARS) is a direct observation scale which has been validated for use in young children (Puhl et al. 1990). While the CARS method is time-consuming and resource intensive, and therefore not suitable for large population-based studies, it has been widely used in methodological studies of young children (O'Hara et al. 1989; Oliver et al. 2011; Sirard et al. 2005; Van Cauwenberghe et al. 2011).

The present study therefore aimed to determine the accuracy of measurement of time spent in sedentary behaviour, LPA and MVPA upon the application of cut-point thresholds developed specifically for pre-school children by Pate et al. (2006), Van Cauwenberghe et al. (2011), Sirard et al. (2005), and Reilly et al. (2003). The cut-points used by Puyau et al. (2002), which are based on a calibration study of older children but which have been used in studies of pre-school children (Fisher et al. 2005b), were also applied. Finally, the cut-point of $\leq$ 100 cpm developed from calibration studies which have used energy expenditure (1.0 - 1.5 METs) for sedentary behaviour was also examined (Pate et al. 2011; Treuth et al. 2004a). The sedentary cut-point of $\leq$ 100 cpm is frequently cited as being appropriate to classify sedentary behaviour (Evenson et al. 2008; Pate et al. 2011; Trost et al. 2011) and has been calibrated in school-aged children (5 to 9 y) by Evenson et al. (2008). Comparison with the direct observation criterion method, using the Children's Activity Rating Scale (CARS) (Puhl et al. 1990), was used to determine accuracy.

The aim of this study was:

- To validate Actigraph accelerometer cut-points for estimating physical activity and sedentary behaviour in pre-school children during free-play.

A secondary aim of this study was:

- To compare accelerometry cut-points for sedentary behaviour, light intensity, and MVPA for estimates of physical activity behaviour in pre-school children during free-play.

6.2 METHODS

Data were collected from a convenience sample of 31 children aged 3 to 5 years recruited from pre-schools in Edinburgh (15 males, 16 females, mean (SD) age: 4.3 (0.8) y, height: 104.8 (6.3) cm, weight: 17.7 (2.5) kg, BMI 16.1 (1.1) kg/m^2). The mean BMI *z*-score was 0.20, with 90% classified as 'healthy' weight and 10% as overweight/obese i.e. BMI at or above 85th centile relative to UK population reference data (Cole 2002). Children were video recorded while they engaged in 1 hour of free-play during their usual outdoor play-time in the nursery setting. Each child wore a GT1M accelerometer on an elasticated belt around their waist. The accelerometers were set to record data in 1-s epochs. The details of the methods are outlined in Chapter 2.

Data from the accelerometers were transferred to an Excel spreadsheet and the data from the 1-s epochs were reintegrated into 15-s epochs to allow for comparison with the CARS data (Edwardson and Gorely 2010). Using a programme developed with Visual Basic for Applications the data were processed within Excel using the cut-points for sedentary, LPA and MVPA intensity levels as defined by Puyau et al. (2002) (Sed^{pu}, LPA^{pu} MVPA^{pu}), Van Cauwenberghe et al. (2011) (Sed^{va}, LPA^{va}, MVPA^{va}) and by Sirard et al. (2005) (Sed^{s}, LPA^{s}, MVPA^{s}) (Table 6.1). In addition, the cut-points for MVPA by Pate et al. (2006) (MVPA^{pa}) and for sedentary behaviour by Reilly et al. (2003) (Sed^{r}) were applied. The cut-point of $\leq$ 100 cpm (Sed^{ev}) for sedentary behaviour was also examined, as this cut-point is cited as being appropriate for sedentary behaviour (Evenson et al. 2008; Pate et al. 2011; Treuth et al. 2004a), and has been calibrated in children (5 - 8 y) by Evenson et al. (2008). The cut-points were divided to allow analysis of 15-s epochs, as undertaken in earlier studies (Nilsson et al. 2002). The 15-s epoch count thresholds used in this study are listed in Table 6.1.

Table 6.1: Summary of the published Actigraph cut-points for young children.

Accelerometer model	Authors	No. participants (n), age range, mean (SD)	Criterion measure		Counts/15 s		
					Sed	LPA	MVPA
7164	Evenson et al. (2008)	33, 5 - 8 y	VO_2 (portable metabolic unit)		≤ 25	≥ 26 to 573	≥ 574
7164	Pate et al. (2006)	29, 3 - 5 y, 4.4 (0.8) y	VO_2 (portable metabolic unit)		NA	NA	≥ 420
7164	Puyau et al. (2002)	26, 6 - 16 y, 10.7 (2.9) y	VO_2 (whole room calorimetry)		< 200	≥ 200 to < 800	≥ 800
7164	Reilly et al. (2003)	30, 3 - 4 y, 3.7 (0.5) y	CPAF		< 275	NA	NA
7164	Sirard et al. (2005)	16, 3 - 5 y	CARS	3 y olds 4 y olds 5 y olds	< 301 < 363 < 398	≥ 302 to 614 ≥ 364 to 811 ≥ 399 to 890	≥ 615 ≥ 812 ≥ 891
GT1M	Van Cauwenberghe et al. (2011)	18, 4 - 6 y, 5.8 (0.3) y	CARS		< 373	≥ 373 to < 585	≥ 585

CARS: Children's Activity Rating Scale; CPAF: Children's Physical Activity Form; LPA: light physical activity; MVPA: moderate-to-vigorous physical activity; NA: not available; Sed: sedentary behaviour.

Data were imported into SPSS (version 17) for analysis. Normality tests were conducted using the Shapiro-Wilks statistic as the sample was less than 50. All data on minutes spent in different intensities, except for time spent in LPA when the Van Cauwenberghe et al. (2011) and the Sirard et al. (2005) cut-points were applied, were found to be not normally distributed (Appendix VI; Table VI: a). As a result, the median and interquartile range (IQR) values are reported. The median (IQR) number of minutes of sedentary behaviour, LPA and MVPA were calculated. Using the Friedman's Repeated Measures ANOVA, the difference between the number of minutes of sedentary behaviour, LPA and MVPA were calculated from each of the cut-points and the CARS at 15-s epochs. Post hoc analysis was undertaken using the Wilcoxon signed rank test, the non-parametric equivalent of the paired *t*-test. To reduce type I error, a Bonferroni correction was applied; the significance level was set at $p < 0.01$.

To assess the accuracy of the different accelerometry cut-points as an absolute measure of physical activity, comparison was made between the number of minutes of sedentary behaviour, LPA and MVPA as estimated by the different cut-points and then compared with the estimates as determined by the criterion measure of direct observation. The Bland and Altman approach was used to examine the relationship between the minutes estimated using the cut-points and the estimates using the criterion measure (Bland and Altman 1986).

6.3 RESULTS

The median (IQR) number of minutes of CARS data collected were 46 (16.0) min during the 1 hour of direct observation. This equates to the scoring of 5704, 15-s epoch observations for the 31 participants.

Table 6.2 presents a summary of the physical activity levels of the sample during the period of direct observation while at pre-school. Results of the Friedman's ANOVA revealed that there was a significant difference in time spent in sedentary behaviour (χ^2 (5) = 134.1, $p < 0.05$), LPA (χ^2 (3) = 81.5, $p < 0.05$) and MVPA (χ^2 (4) = 94.8, $p < 0.00$) as estimated by each of the different cut-points and the CARS

Table 6.2: Median (IQR) minutes of sedentary behaviour, light physical activity and MVPA for different cut-points and the CARS.

	Time (min)		
	Sed	**LPA**	**MVPA**
Puhl et al. (1990): CARS	24.0 (16.8)	17.0 (7.0)	4.8 (6.0)
Pate et al. (2006)	-	-	14.3 (11.8)*
Van Cauwenberghe et al. (2011)	30.0 (15.0)*	7.3 (3.5)*	8.8(9.3)*
Sirard et al. (2005)	29.8 (14.5)*	12.5 (7.3)*	5.5 (6.0)
Puyau et al. (2002)	21.3 (10.0)	19.8 (9.0)	4.3 (6.0)
Reilly et al. (2003)	25.0 (11.3)	-	-
Evenson et al. (2008)	7.8 (5.8)*	-	-

*CARS: Children's Activity Rating Scale (criterion measure); LPA: light physical activity; MVPA: moderate-to-vigorous physical activity; Sed: Sedentary behaviour. *indicates those values which are significantly different from the CARS p < 0.01 with Bonferroni correction.*

Post-hoc analysis using the Wilcoxon test revealed that there was no significant difference for estimates of time spent in Sed^{pu} (*Mdn* = 21.3) and the CARS (*Mdn* = 24.0), z = -0.42, p = 0.7, r = 0.05 , and for the Sed^{r} (*Mdn* = 25.0) and the CARS (*Mdn* = 24.0, z = -2.5, p = 0.01, r = -0.3). No significant difference was found for estimates of time spent in LPA^{pu} (*Mdn* = 19.3) and the CARS (*Mdn* = 17.0, z = 2.3, p = 0.02, r = -0.3). For MVPA there was no significant difference between the $MVPA^{pu}$ (*Mdn* = 4.3) estimate and the CARS (*Mdn* = 4.8, z = -1.9, p = 0.05, r = 0.2) and the $MVPA^{s}$ (*Mdn* = 5.5) estimate and the CARS (*Mdn* = 4.8, z = -1.3, p = 0.2, r = -0.2). As the p values for the Sed^{r}, and LPA^{pu} estimates are close to the predetermined level of significance, (e.g. $p < 0.01$ and $p < 0.02$ respectively), these results could suggest a trend in the data, despite their not being significantly different from the CARS. In addition, there is a medium effect size r = - 0.3 for these variables. There was a significant difference between the CARS and the estimations made by the remaining cut-points, and the levels of significance with the effect sizes are provided in Appendix VI (Appendix Table VI.i).

Bland and Altman plots were undertaken to examine the estimated number of minutes of sedentary behaviour, LPA and MVPA produced by the different cut-points and the CARS criterion measure. A summary of the Bland and Altman output is presented in Table 6.3. Plots presented are for the Sirard et al. (2005), Puyau et al. (2002), Van Cauwenberghe et al. (2011), and Reilly et al. (2003) cut-points for sedentary behaviour and for the Sirard et al.

(2005), Puyau et al. (2002) and Van Cauwenberghe et al. (2011) cut-points for LPA and MVPA (Figure 6.2 and Figure 6.3).

The results suggested a negative bias when the Sed^{s}, Sed^{va} and Sed^{r} cut-points were compared with the CARS, with the cut-points resulting in a greater estimation of time spent in sedentary behaviour than the CARS. While the Sed^{ev} cut-points suggested a large positive bias in relation to the CARS, this cut-point resulted in an underestimation of time spent in sedentary behaviour in relation to the CARS. Application of the Puyau et al. (2002) cut-points for sedentary behaviour resulted in a mean difference with the CARS that was close to zero minutes (mean difference (LOA): 0.4 (-13.3 to 14.0) min), however there were wide limits of agreement. For LPA, the LPA^{va} and LPA^{s} cut-points had a lower estimation of time spent in LPA compared to the CARS. Finally, there was a positive bias, with an underestimation in MVPA, between the $MVPA^{s}$ and the $MVPA^{pu}$ cut-points and the CARS, and a negative bias, with an overestimation in MVPA time, for the $MVPA^{va}$ and the $MVPA^{pa}$ cut-points and the CARS.

Table 6.3: Bland and Altman output for the mean difference (LOA) between cut-point estimates and the CARS.

	Time (min)	
	dm	***LOA***
Sed^{ev} versus CARS	13.2	-2.2 to 28.6
Sed^{pu} versus CARS	0.4	-13.3 to 14.0
Sed^{r} versus CARS	-3.6	-17.6 to 10.4
Sed^{s} versus CARS	-7.2	-20.2 to 5.7
Sed^{va} versus CARS	-8.2	-22.7 to 6.2
LPA^{pu} versus CARS	-2.2	-15.0 to 10.5
LPA^{s} versus CARS	6.3	-6.0 to 18.6
LPA^{va} versus CARS	11.4	-1.4 to 24.2
$MVPA^{s}$ versus CARS	0.8	-6.2 to 7.8
$MVPA^{pa}$ versus CARS	-8.7	-19.9 to 2.5
$MVPA^{pu}$ versus CARS	1.7	-8.0 to 11.5
$MVPA^{va}$ versus CARS	-3.3	-11.8 to 5.3

LOA: limits of agreement; dm: mean difference; CARS: Children's Activity Rating Scale; ev: Evenson et al. (2008) *cut-points; MVPA: moderate-to-vigorous activity; pa: Pate et al.* (2006) *cut-points; pu: Puyau et al.* (2002) *cut-points; r: Reilly et al.* (2003) *cut-points; s: Sirard et al.*(2005) *cut-points; va: van Cauwenberghe et al.* (2011) *cut-points.*

Figure 6.1: Bland and Altman plots assessing agreement between CARS and minutes of sedentary time (Sed) estimated with cut-points for: (i) Sirard et al (2005), (ii) Reilly et al (2003), (iii) Van Cauwenberghe et al (2011), (iv) Puyau et al (2002).

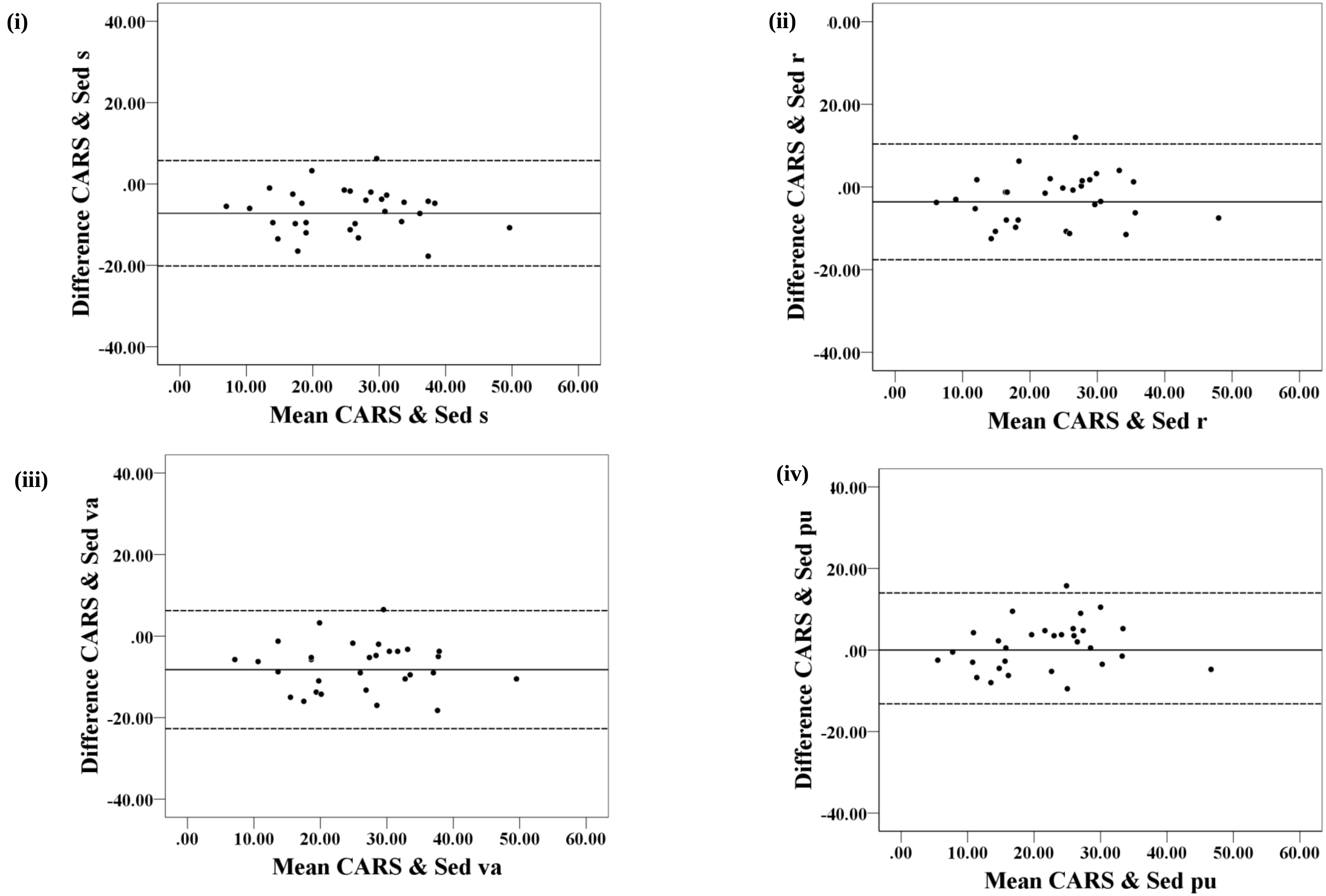
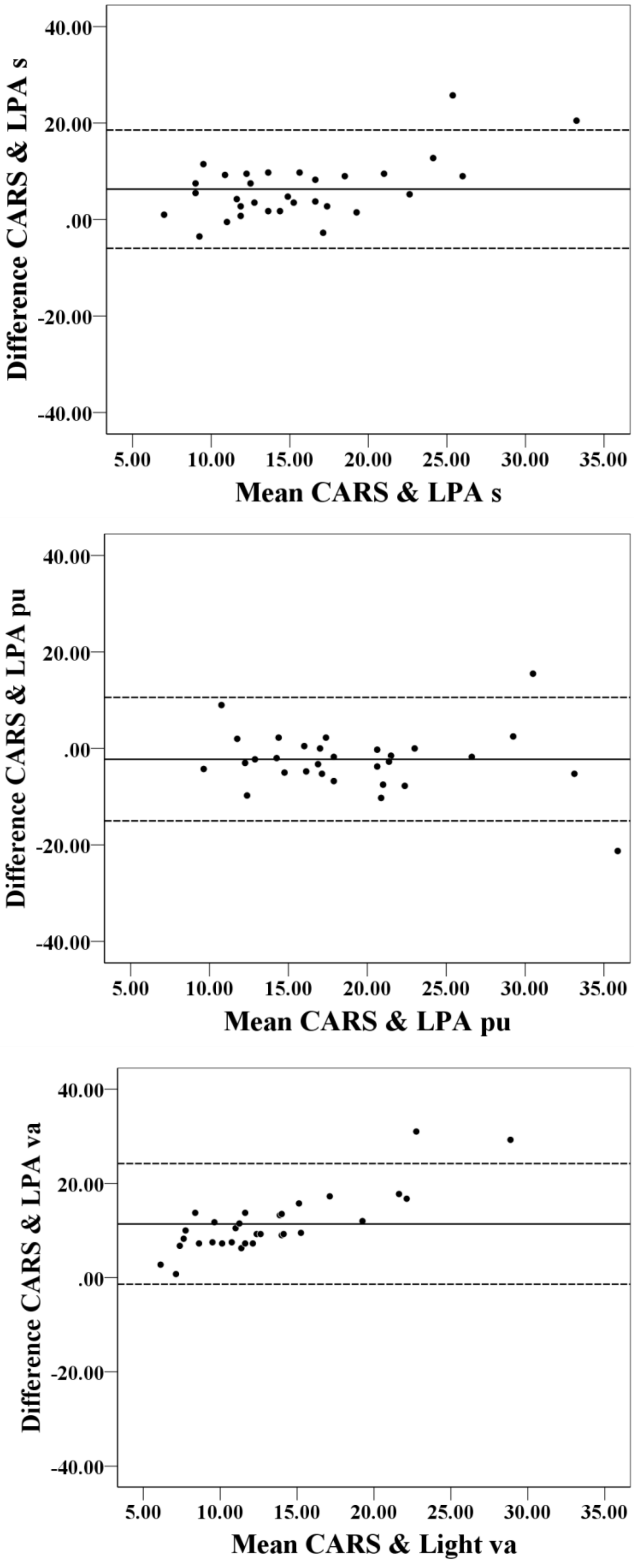


Figure 6.2: Bland and Altman plots assessing agreement between CARS and estimates of minutes of LPA with cut-points for: (i) Sirard et al (2005), (ii) Reilly et al (2003), (iii) Van Cauwenberghe et al (2011), (iv) Puyau et al (2002).

(i)

(ii)

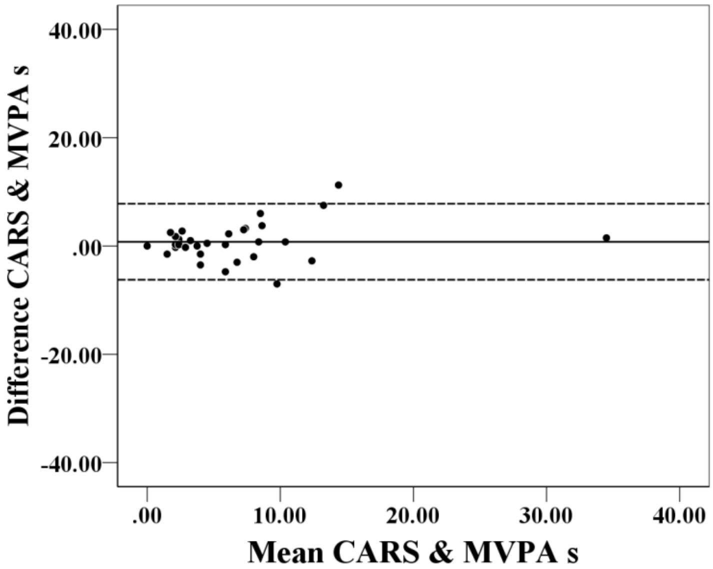
(iii)



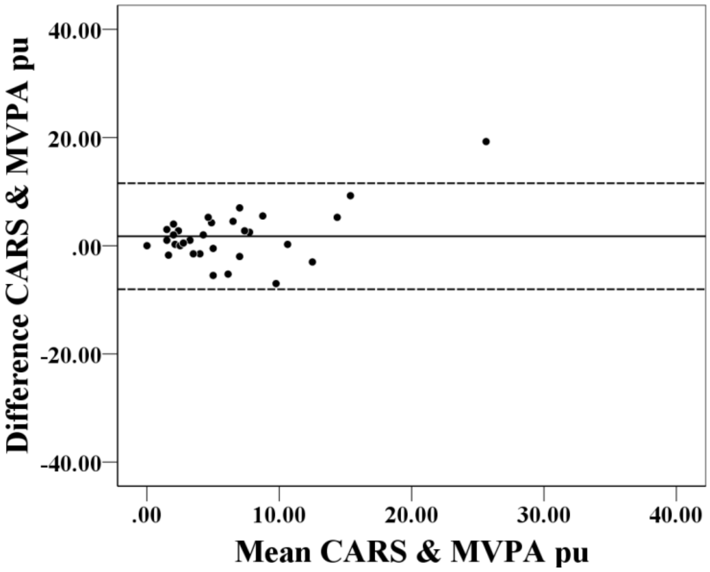
CARS: Children's Activity Rating Scale; LPA: light physical activity.

Figure 6.3: Bland and Altman plots assessing agreement between CARS and estimates of minutes of MVPA with cut-points for: (i) Sirard et al (2005), (ii) Reilly et al (2003), (iii) Van Cauwenberghe et al (2011), (iv) Puyau et al (2002).

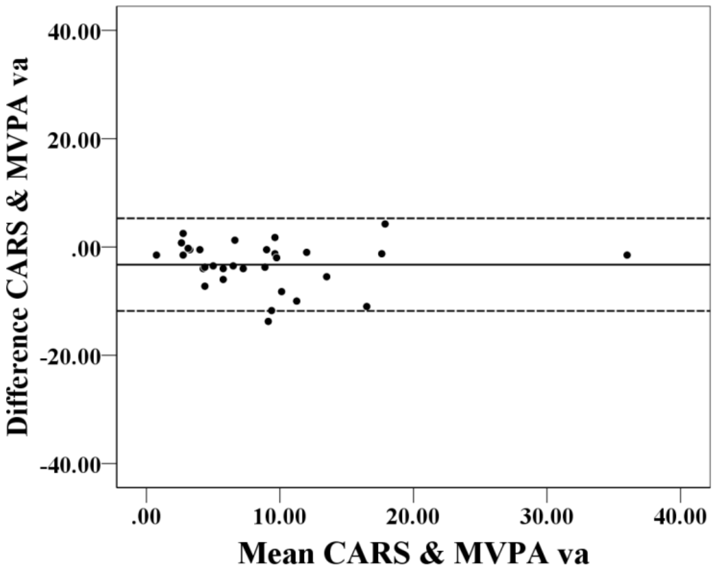
(i)



(ii)



(iii)



CARS: Children's Activity Rating Scale; MVPA: moderate-to-vigorous physical activity.

6.4 DISCUSSION

The results of the current study indicate relatively large errors at the individual level (wide limits of agreement) when accelerometry output was compared with the CARS. On a group level, $MVPA^{pu}$ and $MVPA^{s}$ provided estimates which did not differ significantly compared with the criterion direct observation method. There was a mean difference of 0.8 min for the Sirard et al. (2005) cut-points and 1.7 min for the Puyau et al. (2002) cut-points, with a bias suggesting an overestimation of time spent in MVPA in comparison to the CARS criterion measure. Only the Sed^{r}, Sed^{pu} and LPA^{pu} cut-points provided accurate estimates of sedentary behaviour and LPA (mean difference of -3.6, 0.4 and -2.2 min respectively). However, with the $p = 0.02$ for LPA^{pu} and $p = 0.01$ for Sed^{r} being close to the predetermined level of significance ($p < 0.01$), this could indicate a trend for a difference between the CARS and the estimated time in sedentary behaviour with the Sirard et al. (2005) cut-points and for LPA using the Puyau et al. (2002) cut-points. In addition, there was a medium effect size for both these estimates ($r = -0.3$). To interpret the meaning of these findings, the 95% confidence interval (CI) of the effect sizes (*ES*) were calculated using Equation 6.1.

Equation 6.1: Calculation of confidence intervals (CI)

$$\mathrm{CI} = ES - 1.96 \pm se$$

(Nakagawa and Cuthill 2007)

The *se*, which is the asymptomatic standard error of the effect size (Ivarsson et al. 2013) can be calculated using Equation 6.2.

Equation 6.2: Calculation of standard error of the effect size

$$se(r) = \frac{1}{\sqrt{n-3}}$$

(Ivarsson et al. 2013)

The results suggest that the 95% confidence interval for LPA^{pu} was from -0.04 to -0.6. and for $Sirard^{s}$ this was -0.06 to -0.6. Confidence intervals are used to describe where most (in this case 95%) of the sample are located (Thompson 2002). As the confidence intervals for the effect sizes do not include zero, it is likely that some effect has taken place (Ivarsson et al. 2013). In contrast, $MVPA^{pu}$, with the effect size 0.2 and 95% confidence interval of -0.05 to 0.01, does include zero, which could indicate that there is indeed no difference between the CARS and the $MVPA^{pu}$ estimate, despite $p = 0.05$ being close to the predetermined level of significance set at $p < 0.01$. However these results should be viewed cautiously as confidence intervals are sensitive to violations of normality (Thompson 2002), which could lead to inaccuracies in results (Ivarsson et al. 2013).

Although not presented in the results it is interesting to note that combining the LPA^{pu} and the $MVPA^{pu}$ resulted in a median (IQR) of 24.5 (12.0) min of TPA, which was not significantly different ($z = -0.5$, $p = 0.7$, $r = -0.06$) from the CARS estimate for TPA (median (IQR) 23.0 (13.5) min). The median (IQR) min of TPA using the Sirard et al. (2005) cut-points was 16.8 (10.8) min and for the Van Cauwenberghe et al. (2011) cut-points this was 16.0 (11.8) min. Both estimates of minutes of TPA from the application of the Sirard et al. (2005) and Van Cauwenberghe et al. (2011) cut-points were significantly different from the CARS criterion measure ($z = -4.4$, $p = 0.00$, $r = -0.6$; $z = -4.5$ and $p = 0.00$, $r = -0.6$ respectively). The agreement with the estimated minutes of TPA with Puyau et al. (2002) cut-points and the criterion measure is important given the recent changes in recommendations for pre-school children which have been expressed in terms of 'total' physical activity (combining LPA and MVPA) (Australian Government, Department of Health and Ageing 2009; Canadian Society of Exercise Physiology 2012; Department of Health, Physical Activity, Health Improvement and Protection 2011).

Use of the $MVPA^{pa}$ cut-point produced a statistically significant difference and large biases relative to direct observation, with overestimation of time spent in MVPA. Application of the $MVPA^{pu}$, $MVPA^{pa}$, $MVPA^{va}$ and $MVPA^{s}$ cut-point to the different age groups in the sample (3 year olds (n= 10), 4 year olds (n = 13), and 5 year olds (n = 8)) revealed that the Sirard et al. (2005) age-specific cut-points for the 3 year olds in this study were the most accurate (0.7 min difference from the criterion) and the Puyau et al. (2002) cut-points were the most accurate for the 5 year olds (0.9 min difference from the criterion). One of the outliers in the $MVPA^{pu}$ Bland and Altman plot (Figure 6.3) is a 3-year-old child for whom the higher Puyau et al. (2002) cut-points for MVPA resulted in a greater mean difference, i.e. an underestimation of time spent in MVPA in relation to the CARS, while the Sirard et al. (2005) cut-points lead to a more accurate categorisation of intensity level for this child. It is, however, recognised that these subsamples are small and these results should be interpreted with caution.

Sirard et al. (2005), Pate et al. (2006) and Van Cauwenberghe et al. (2011) have identified cut-point thresholds for Actigraph accelerometers specific to the pre-school population. One explanation for the differences between estimates of physical activity derived from these cut-points and the CARS measure in the current study may be due to the differing calibration methods used in studies to develop cut-points. Pate et al. (2006) used indirect calorimetry to calibrate their accelerometry cut-points, which is a criterion measure of energy expenditure. In contrast, Sirard et al. (2005) and Van Cauwenberghe et al. (2011) used the CARS direct observation scale, which is a behavioural criterion measure. Freedson et al. (2005) has suggested that behavioural approaches to calibration are particularly useful in studies of young children, where measurement and interpretation of energy expenditure can be difficult. Behavioural methods also avoid errors associated with extrapolation from treadmill activity to free-living behaviours. There are, however, also limitations with direct observation methods, such as risk of reactivity from participants and problems with the accuracy of activity classification, as the rating of intensity by an observer is subjective (Westerterp 2009).

The current study may also be limited as it relied on the CARS direct observation scale as the criterion measure, where ideally this could have been combined with either VO_2 or energy expenditure to gain a more comprehensive view of physical activity. Interestingly, however, Puyau et al. (2004) also used calorimetry to calibrate their accelerometry cut-points and the findings of the current study suggest that no significant bias exists between the number of minutes of MVPA with the Puyau et al. (2002) cut-points and the CARS.

It is important to note that with the exception of the Van Cauwenberghe et al. (2011), all the other cut-points were calibrated for the earlier generation of Actigraph, the 7164 model, which has a different internal technology to the GT1M (John and Freedson 2012). This has implications for the validity of the 7164 cut-points for the GT1M and their suitability for processing the GT1M accelerometry data is unclear. Comparability between the different generations of Actigraph has not been formally established in the pre-school population. The next chapter focuses on comparison between the different generations of Actigraph.

In conclusion, marked differences between cut-points and the impact that these differences have on apparent levels of MVPA and time spent in sedentary behaviour have been highlighted in a number of earlier studies (Guinhouya et al. 2006; Mota et al. 2007; Pate et al. 2011; Reilly et al. 2008). However, while these studies have described the differences in outcome with different cut-points, in contrast to the current study, they have not addressed the issue of accuracy of various cut-points.

The present study indicates that the Pate et al. (2006) cut-point and the Van Cauwenberghe et al. (2011) cut-point significantly overestimated minutes of MVPA in pre-school children compared to direct observation. Use of the Puyau et al. (2002) and Sirard et al. (2005) cut-points produced estimates of MVPA which were not significantly different from the criterion measure at a group level. Importantly, the application of the Puyau et al. (2002) thresholds for sedentary, LPA and for TPA, provided the only estimates that were not significantly different from the criterion method providing agreement at a group level. However, even with the 'best' cut-points, there were large errors as revealed by the Bland and Altman plots which limit the use of cut-points to accurately measure physical activity intensity and sedentary behaviour at an individual level.

CHAPTER 7 : COMPARISON BETWEEN DIFFERENT GENERATIONS OF ACTIGRAPH ACCELEROMETER DURING MECHANICAL CALIBRATION AND DURING FREE-PLAY WITH PRE-SCHOOL CHILDREN

7.1 INTRODUCTION

The Actigraph accelerometer is one of the most widely utilised motion sensors in children's physical activity research (Trost et al. 2011), and the uniaxial Actigraph accelerometers have been validated for use with pre-school children (Pate et al. 2006). The Actigraph 7164 has been used in the majority of calibration studies with pre-school children to establish cut-point thresholds, which are used to convert the raw accelerometry count output per epoch into time spent at different intensities (Pate et al. 2006; Reilly et al. 2003; Sirard et al. 2005). However, the 7164 has been discontinued and was in 2005 replaced by the uniaxial GT1M and subsequently by the triaxial GT3X and the GT3X+ models. The GT1M and the GT3X+ have a different internal technology from the 7164 accelerometer. These newer models use a Micro-Electro-Mechanical System (MEMS) solid state accelerometer instead of the piezoelectric cantilevered beam accelerometer used in the 7164 model (John et al. 2010). In addition, while both accelerometers capture time-varying acceleration in the range of 0.05 to 2 *G*, there are differences in the filtering and sampling frequencies between models. As a result, the same acceleration could result in different count output (Chen et al. 2012) and the two types of accelerometer cannot necessarily be used interchangeably.

To date, comparison studies of the GT1M and 7164 models have focused on adolescent and adult populations, or have made use of a mechanical set-up to compare accelerometry output. Results from these comparison studies have either reported no significant difference between each model's activity counts during treadmill running and walking with adults (John et al. 2010) or significant differences using a mechanical set up (Kozey et al. 2010b; Rothney et al. 2008). One study found the GT1M counts to be on average 9% lower than those derived from the 7164 model in adolescents, recommending the application of a correction factor to the GT1M data (7164=GT1M/0.91) (Corder et al. 2007). While recent research suggests good agreement between the GT1M and GT3X models (Kaminsky and Ozemek 2012; Robusto and Trost 2012; Vanhelst et al. 2012), comparability of these models with the older 7164 model has not been established in child populations or examined with children under free-living conditions. This is important, as it is unclear whether the cut-points calibrated for the 7164 are valid for the GT1M and GT3X Actigraph models. Despite this, the cut-points are used between the models interchangeably. Given that several large longitudinal studies are reporting on the tracking of physical activity in children (Kristensen et al. 2008; Kwon and Janz 2012; Metcalf et al. 2011; Mitchell et al. 2012), it is important to ensure that systematic errors have not taken place as a consequence of researchers using different generations of accelerometers, or applying cut-points developed for the 7164 model to the later Actigraph models. Recent research has identified the need for more studies comparing different generations of accelerometers under free-living conditions on different populations (Ried-Larsen et al. 2012). The current study therefore aimed to compare the GT1M and the 7164 Actigraph models in pre-school children under free-living conditions.

In summary, while a small number of studies have compared Actigraph models, there is a lack of consensus as to whether there are differences in output between the 7164 and the GT1M and GT3X models. It is also unclear whether the application of a correction factor to the GT1M output is necessary to make it comparable to the output from the older 7164 model (upon which many accelerometry thresholds are based). Importantly, whether any differences in raw counts between models will impact on time spent at different intensities of physical activity for pre-school children has not been investigated. Finally, given the recent evidence-based recommendations to encourage children to increase their total daily physical activity to 180-minutes per day (Australian Government, Department of Health and Ageing 2009; Canadian Society of Exercise Physiology 2012; Department of Health, Physical Activity, Health Improvement and Protection 2011), it would also be important to establish comparability in terms of time spent in TPA, to allow cross-comparison between studies.

The current study proposed to bring new information on whether there will be any meaningful differences between time spent in sedentary behaviour, LPA, MVPA and in TPA between Actigraph models.

The aims of this study were:

- To compare different generations of Actigraph accelerometer during mechanical calibration (study 1: In vitro comparison).
- To compare different generations of Actigraph accelerometer in pre-school children during 1 hour of free-play (study 2: In vivo comparison).

7.2 METHODS

7.2.1 Study 1: In vitro comparison between 7164 and GT1M model using mechanical calibration

A mechanical calibration set up, based on a model by Brage et al. (2003), was developed and consisted of three wheels connected by a horizontally placed linkage with a radius setting of 25 mm. Six accelerometers could be positioned along their χ axis on the set-up. Rotation of the wheels caused the accelerometers to move through sinusoidal acceleration in two axes (anteroposterior and mediolateral axes). The accelerometers were set to collect data in 1-s epochs. To test for intra-instrument agreement six accelerometers of each model were tested three times at different frequencies of rotation: 1 Hz, 2 Hz and 3 Hz. To calculate the average acceleration (ā) produced by the mechanical set up at each frequency the equation by Brage et al. (2003) (Equation 7.1), where acceleration is a function of radius (r) and frequency (f) of oscillation was used.

Equation 7.1: Calculation of average acceleration

$$\bar{a} = 8 \cdot \pi \cdot r \cdot f^2$$

(Brage et al. 2003)

The accelerometers were therefore subjected to average accelerations of 0.63, 2.51 and 5.65 $m \cdot s^{-2}$ respectively. Each test lasted 5 minutes. To test for inter-instrument agreement between the GT1M and the 7164, two accelerometers of each model were tested by simultaneously subjecting them to 15 accelerations between 0.5 to 3.3 Hz in average increments of 0.2 Hz. This corresponds to average accelerations ranging from 0.16 to 6.84 $m \cdot s^{-2}$. Each test lasted five minutes, at each of the accelerations.

Data from the mechanical calibration process were imported to SPSS (version 17.0) for analysis. The first and last minutes of acceleration data were excluded so that the middle 3 minutes of data were used. To determine the intra-unit variability, the standard error of measurement (*SEM*) and coefficient of variation (CV) were calculated. To determine inter-model variability an interclass correlation coefficient (ICC) was calculated. The mean count per minute (cpm) output for accelerometers of each model was plotted to view how they responded during the 15 different accelerations.

7.2.2 Study 2: In vivo comparison between the 7164 and GT1M model during free-play

Twenty-three healthy children were recruited from two pre-schools in Edinburgh Scotland (10 boys, 13 girls, mean (SD) age: 4.3 (0.8) y, height: 106.2 (6.2) cm, weight: 18.3 (2.5) kg, BMI 16.2 (1.1) kg/m^2). The mean BMI *z*-score was 0.20, with 87% classified as 'healthy weight' and 13% classified as overweight or obese i.e. BMI at or above the 85^{th} centile relative to the UK reference data (Cole 2002)). Children were video-recorded while they engaged in 1 hour of unstructured free-play in the pre-school setting during the children's usual timetabled outdoor playtime. The GT1M and 7164 were placed adjacent to each other, on an elasticised belt around the child's waist, with one monitor posterior to the other, with a gap of 3 cm. The order of monitor placement was alternated between subjects. The accelerometers were synchronised for data collection with the laptop PC, which was also synchronised with the video-recording equipment; both accelerometers were set to collect data in 1-s epochs. Details of the procedures are outlined in the general methods chapter (Chapter 2).

As this was a paired study design to enable comparison between models, it was expected to have a high power. At initial analysis with a sample of 23 subjects, significant differences between models were observed and therefore it was felt to be sufficiently powered at this point and no further recruitment was undertaken to increase the sample size.

Video data were coded using the CARS direct observation scale (Puhl et al. 1990); the details of the process for coding are outlined in the general methods chapter (Chapter 2). Following data collection, data from the CARS and the accelerometers were transferred to Excel and imported to SPSS (version 17). Data from the accelerometers were reintegrated from 1-s epochs into 15-s epochs to allow comparison with the CARS data, which had been coded in 15-s time periods.

The mean cpm was calculated for each participant for the GT1M and the 7164 monitors. The cpm for each accelerometer model was tested for normality across the sample using the Shapiro-Wilks test and the data were found to be not normally distributed (Appendix VII, Appendix Table VII.i). The median and interquartile ranges (IQR) were reported and agreement between monitors was assessed using Spearman's correlation. Inter-monitor agreement for cpm and limits of agreement (LOA) were evaluated using a Bland and Altman plot (Bland and Altman 1986). This analysis was repeated following the application of a correction factor of 0.91 to the GT1M data. This correction factor has been recommended by Corder et al. (2007) who, in a study of free-living activity of 30 Indian adolescents (mean (SD) age: 15.8 (0.6) y) over a seven day period, reported that the output from the GT1M was on average 9% lower than the 7164. The 1-s epoch data from the accelerometers were converted to 15-s epochs, as undertaken in earlier studies (Nilsson et al. 2002), and two sets of cut-points developed for the 7164 were applied: the age specific cut-points by Sirard et al. (2005) (Table 7.1) and those developed for slightly older children by Puyau et al. (2002). Earlier studies in this thesis suggested that the Puyau et al. (2002) cut-points were accurate at a group level against a criterion measure for estimates of sedentary, LPA and TPA with the GT1M accelerometer. Application of the cut-points allowed for calculation of the number of minutes spent in sedentary behaviour, LPA and MVPA for each model of accelerometer.

Table 7.1: Cut-point thresholds.

Authors	Age	Sed	Counts/15 s LPA	MVPA
Sirard et al. (2005)	3 y olds	< 301	≥ 302 to ≤ 614	≥ 615
	4 y olds	< 363	≥ 364 to ≤ 811	≥ 812
	5 y olds	< 398	≥ 399 to ≤ 890	≥ 891
Puyau et al. (2002)		< 200	≥ 200 to < 800	≥ 800

LPA: light physical activity; MVPA: moderate-to-vigorous physical activity; Sed: sedentary behaviour.

The time spent in TPA was also calculated using the threshold for LPA as the lower cut-point threshold e.g. all activity ≥ 200 counts/15 s for the Puyau et al. (2002) cut-points. The time spent in each intensity period was also calculated once the correction factor of 0.91 had been applied to the GT1M data ($GT1M^{corr}$ = GT1M/0.91). In addition, Corder et al. (2007) also suggested that the GT1M cut-points should be 10% lower, and therefore in this study the 7164 cut-points were reduced by 10% and applied to the GT1M data. The data were tested for normality using the Shapiro-Wilks test and the results are presented in Appendix VII (Appendix Table VII.i).

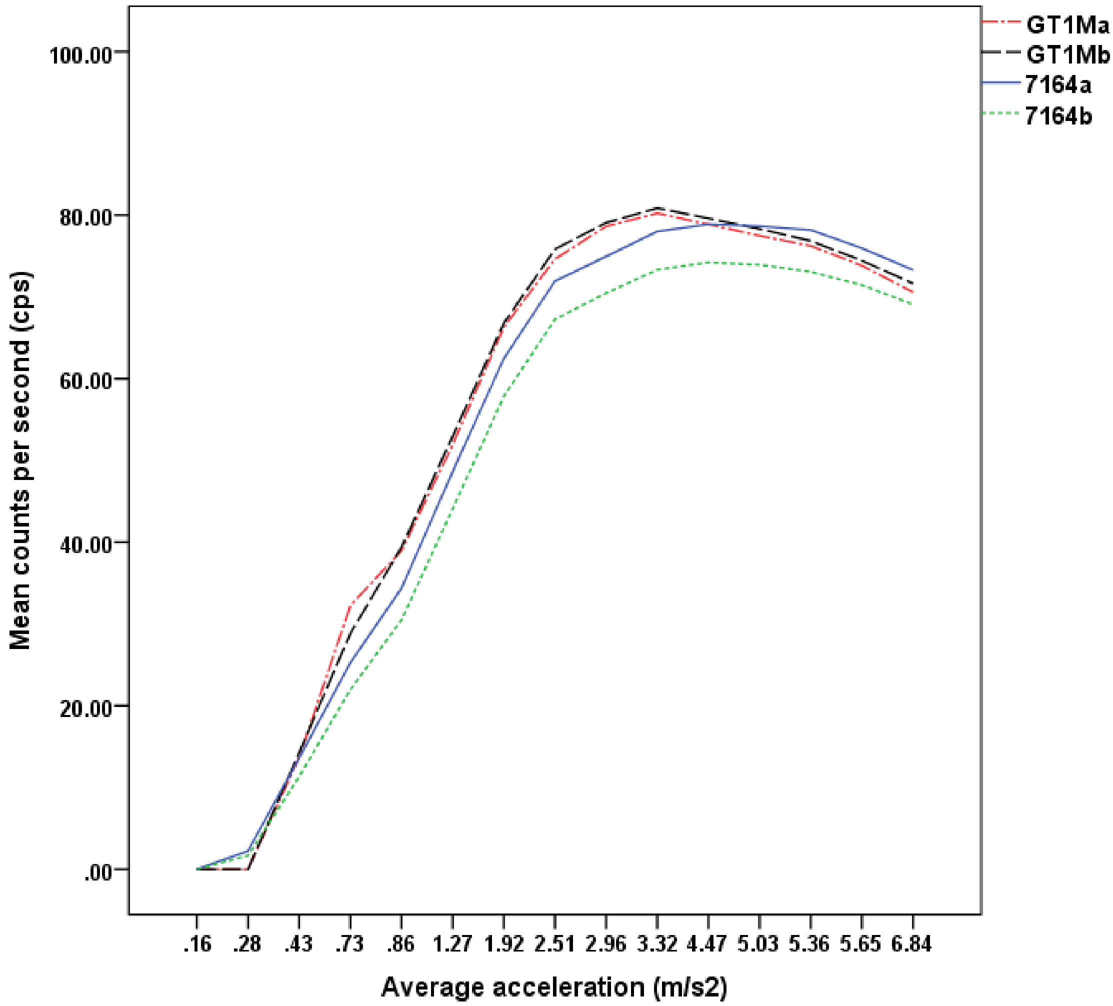
Paired *t*-tests were used to compare differences between percentage time spent at each intensity level between models, and between time spent in TPA and for the corrected GT1M data. To determine the agreement between the accelerometry output against the CARS, the mean differences with limits of agreement (LOA) were calculated (Bland and Altman 1986).

7.3 RESULTS

7.3.1 Study 1: In vitro comparison between 7164 and GT1M model using mechanical calibration

The mechanical calibration revealed that the mean intra-instrument CV for the GT1M was 0.66% and for the 7164 this was 8.28%. The mean *SEM* for the GT1M was 0.62 and for the 7164 the *SEM* was 2.42. The ICC for the GT1M was 0.91 while the 7164 was 0.98. Pearson's correlation coefficient between models was $r = 0.95$, $p < 0.01$ and the mean difference (LOA) were 3.19 (0.16 to 6.22) counts/s. The results of the cps output at different accelerations are plotted in Figure 7.1.

Figure 7.1: Mean counts per second (cps) comparing GT1M and 7164 accelerometers during mechanical calibration at increasing accelerations.



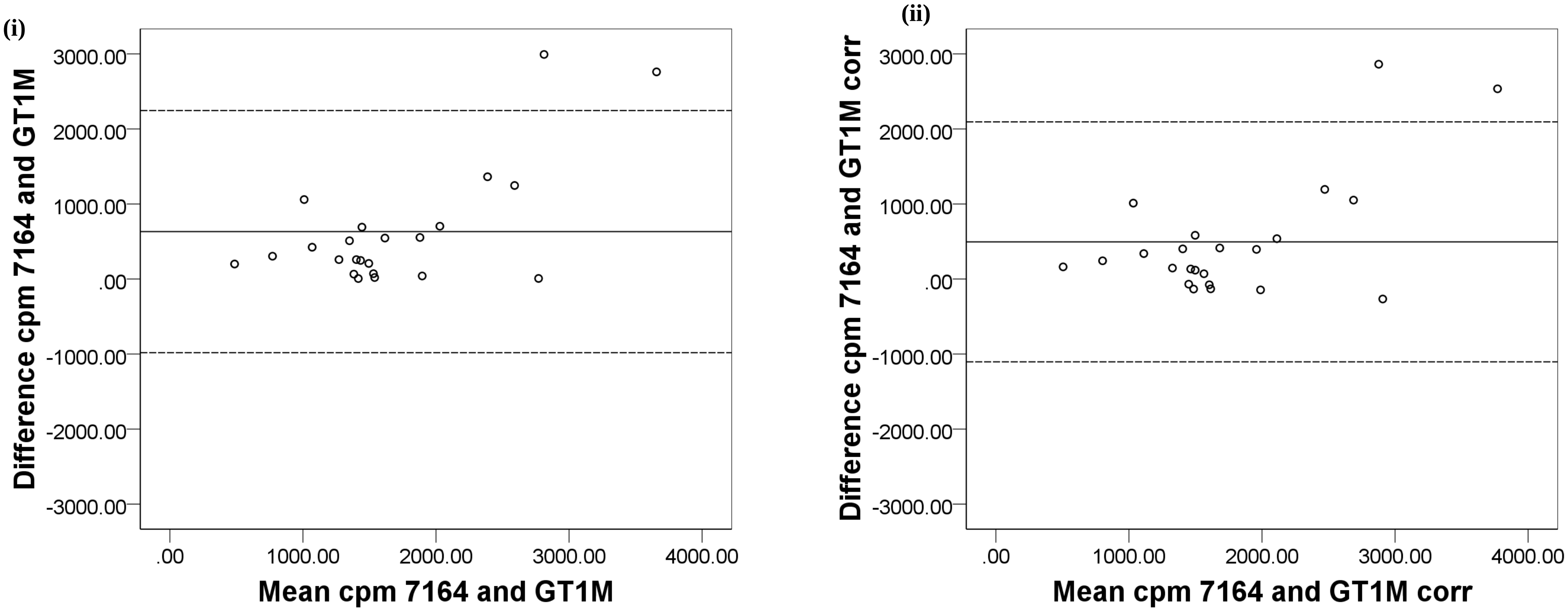
7.3.2 Study 2: In vivo comparison between the 7164 and GT1M model during free-play

During free-play, the median (IQR) cpm for the 7164 was 1599 (962) cpm and for the GT1M this was 1350 (579) cpm. A scatter plot of the mean cpm for each child is presented in Appendix VII (Appendix Figure VII.i).

There was a positive correlation between the cpm for the 7164 and the GT1M accelerometers $r = 0.70$, $p < 0.01$. The mean difference (LOA) between accelerometer models was 632 (-981.7 to 2245.7) cpm. Applying the correction factor of 0.91 to the GT1M data resulted in a mean difference of 495 (-1103.7 to 2092.7) cpm. Figure 7.2 presents the Bland and Altman plot of the mean difference and LOA for the cpm between the 7164 and the GT1M accelerometry data and between the 7164 and the corrected GT1M accelerometry data ($GT1M^{corr}$).

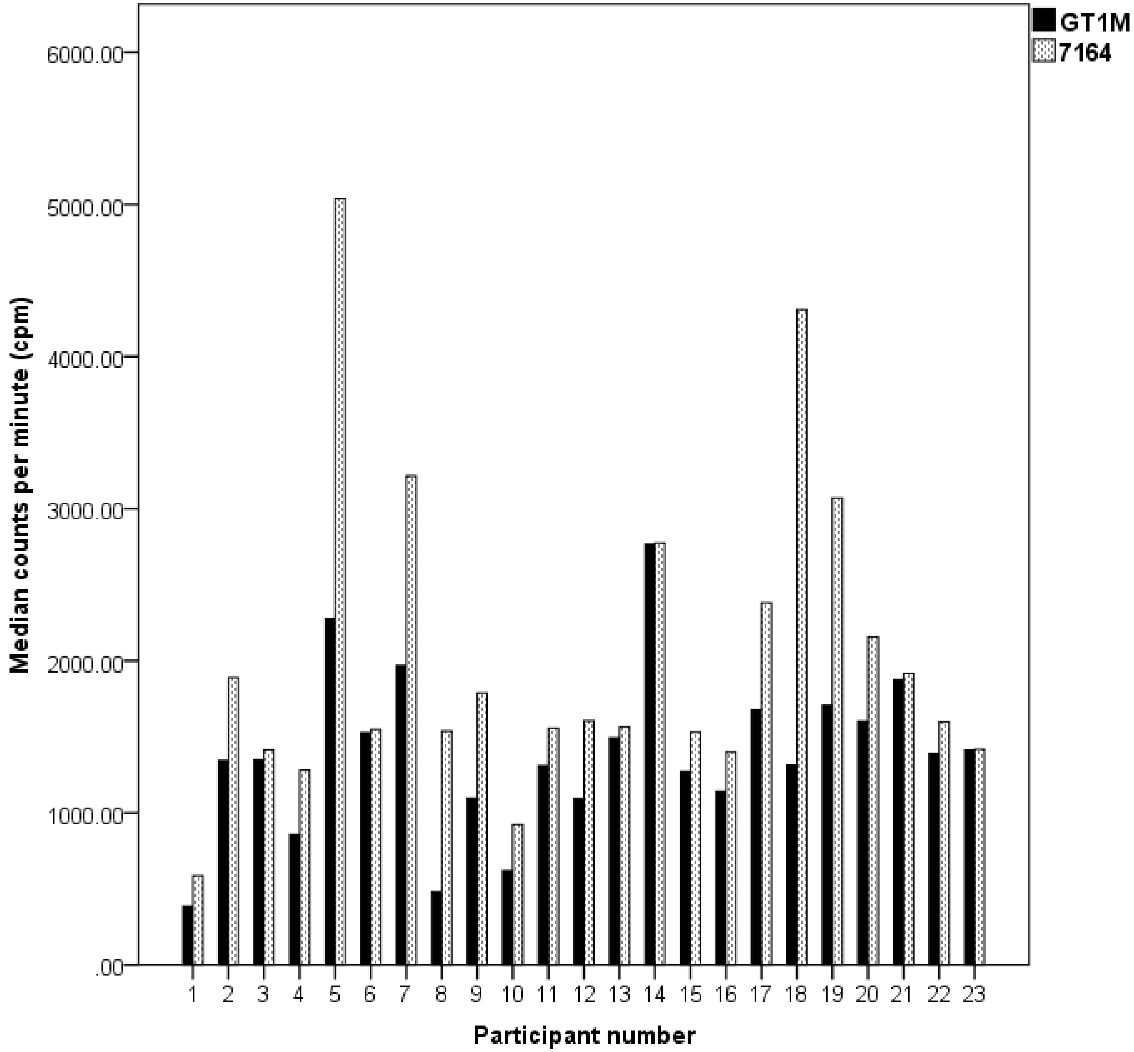
Figure 7.3 presents the median cpm for each participant. It is noticeable that for all participants the median cpm recorded by the 7164 accelerometer was higher than that recorded by the GT1M.

Figure 7.2: Bland and Altman plot of counts per minute (i) between the GT1M and the 7164 (ii) between the GT1M data corrected and 7164 accelerometer.



cpm: counts per minute; GT1M^{corr}: GT1M corrected data (GT1M^{corr}=GT1M/0.91).(i) Mean difference (LOA): 632;(-981.7 to 2245.7) cpm between 7164 and the GT1M. (ii) Mean difference (LOA): 495; (-1103.7 to 2092.7) cpm.

Figure 7.3: Comparison of the median counts per minute between participants (n=23).



The following section will present the results for the time spent in sedentary behaviour and time spent in different intensities, comparing the estimates from the GT1M with those from the 7164 accelerometer. As the data for time spent in sedentary behaviour were normally distributed (Appendix VII; Appendix Table VII: a), the median and SD are reported. Table 7.2 presents the mean (SD) of time spent in sedentary activity using the Sirard et al. (2005) and the Puyau et al. (2002) cut-points. Corrected GT1M data are also presented as $GT1M^{corr10\%}$ where the cut-points have been reduced by 10% and applied to the GT1M data, and the $GT1M^{corr}$ where the GT1M data are altered using the formula $GT1M^{corr} = GT1M/0.91$, as recommended by Corder et al. (2007).

Table 7.2: Mean (SD) minutes spent in sedentary behaviour comparing 7164 against the GT1M with two sets of cut-points.

	Mean min (SD)			
Intensity	**7164**	**GT1M**	**$GT1M^{corr10\%}$**	**$GT1M^{corr}$**
Sed^{s}	27.1 (11.6)	29.0 (12.0)*	**27.6 (11.9)**	31.9 (13.2)*
Sed^{pu}	20.4 (9.5)	**21.6 (10.8)**	10.0 (8.1)*	23.7 (11.8)*

*$GT1M^{corr10\%}$: cut-points reduced by 10%; $GT1M^{corr}$: GT1M/0.9; pu: Puyau et al. (2002) cut-points; Sed: sedentary behaviour; s: Sirard et al. (2005) cut-points. Values shown in bold not significantly different from 7164; * values significantly different from 7164 estimate.*

The results of paired *t*-tests revealed there was a significant difference between the GT1M (M = 29.0, SE = 2.5) estimate and the 7164 estimate (M = 27.1, SE = 2.41, $t(22)$, $p < 0.02$, $r = 0.5$) of time spent in sedentary behaviour when the Sirard et al. (2005) cut-points were applied to the data. Reducing the Sirard et al. (2005) cut-points by 10% for the GT1M data resulted in a non-significant difference (M = 27.6, SE = 2.49, $t(22)$, $p = 0.50$, $r = 0.14$). However, application of a correction factor to the GT1M resulted in a significant difference from the 7164 estimate (M = 31.9, SE = 2.8, $t(22)$, $p < 0.00$, $r = 0.75$). There is a different pattern when the Puyau et al. (2002) cut-points are applied to the GT1M and the 7164 data. This suggests that there was no significant difference between the 7164 estimate (M = 20.4, SE = 2.0) and the GT1M estimate (M = 21.6, SE = 2.2, $p = 1.2$). Reducing the Puyau et al. (2002) cut-points resulted in a significant difference between the GT1M estimate and the 7164 estimate (M = 10.0, SE = 1.69, t(22), $p < 0.00$, $r = 0.84$), as did applying the correction factor to the GT1M data (M = 23.7, SE = 2.5, $p = 0.001$, $r = 0.57$).

The data for time spent in LPA were not normally distributed, therefore the median (IQR) values are reported in Table 7.3.

Table 7.3: Median (IQR) minutes spent in LPA comparing 7164 against the GT1M with two sets of cut-points.

	Median min (IQR)			
Intensity	**7164**	**GT1M**	**GT1M$^{corr10\%}$**	**GT1M^{corr}**
LPA^{s}	9.0 (7.3)	**10.8 (8.3)**	**10.0 (6.3)**	11.8 (9.1)
LPA^{pu}	15.3 (10.8)	18.5 (13.5)*	29.3 (15.5)*	20.3 (14.8)*

*GT1M$^{cor10\%}$: cut-points reduced by 10%; GT1M^{corr}: GT1M/0.91; LPA: light physical activity; pu: Puyau et al. (2002) cut-points; s: Sirard et al. (2005) cut-points. Values shown in bold not significantly different from 7164; * values significantly different from 7164 estimate.*

There was a non-significant difference ($p > 0.01$) for the Sirard et al. (2005) cut-points between the GT1M (*Mdn* = 10.8) and 7164 estimates (*Mdn* = 9.0, $z = -1.7$, $p = 0.08$, $r = -0.25$) and for the cut-points reduced by 10% (*Mdn* = 10.0, $z = -1.75$, $p = 0.08$, $r = -0.26$). There was a significant difference between the 7164 (*Mdn* = 15.3) and the GT1M estimates (*Mdn* = 18.5, $z = -2.39$, $p = 0.02$, $r = -0.35$) when the Puyau et al. (2002) cut-points were applied.

The data for MVPA were not normally distributed and therefore the median (IQR) values are reported and presented in Table 7.4.

Table 7.4: Median (IQR) minutes spent in MVPA comparing 7164 against the GT1M with two sets of cut-points.

	Median min (IQR)			
Intensity	**7164**	**GT1M**	**GT1M$^{corr10\%}$**	**GT1M^{corr}**
MVPA^{s}	6.5 (6.5)	4.3 (6.5)*	5.3 (8.5)*	4.7 (7.1)*
MVPA^{pu}	6.0 (8.5)	3.5 (6.0)*	5.0 (7.0)*	3.8 (6.6)*

*GT1M$^{corr\ 10\%}$: cut-points reduced by 10%. GT1M^{corr}: GT1M/0.91; MVPA: moderate-to-vigorous intensity activity; pu: Puyau et al. (2002) cut-points; s: Sirard et al. (2005) cut-points. * Values significantly different from 7164 estimate.*

All the values for GT1M were significantly different from the 7164, with both sets of cut-points ($p < 0.01$), and application of the correction factor or reducing the cut-points by 10% failed to resolve the difference between the estimates of MVPA.

The data for TPA were normally distributed (Appendix VII, Appendix Table VII.i) and the mean values (SD) are presented in Table 7.5.

Table 7.5: Mean (SD) minutes spent in TPA comparing 7164 against the GT1M with two sets of cut-points.

	Mean min (SD)			
Intensity	**7164**	**GT1M**	**GT1M$^{corr10\%}$**	**GT1M^{corr}**
TPA^{s}	18.4 (9.0)	16.4 (8.5)*	**17.8 (9.6)**	**18.0 (9.3)**
TPA^{pu}	25.0 (11.0)	**23.9 (11.1)**	35.4 (14.1)*	**26.2 (12.2)**

*GT1M$^{corr\ 10\%}$: cut-points reduced by 10%; GT1M^{corr}: GT1M/0.91; pu: Puyau et al. (2002) cut-points; s: Sirard et al. (2005) cut-point; TPA: total physical activity. Values in bold not significantly different from 7164; * values significantly different from 7164 estimate.*

There were significant differences between the TPA^{s} 7164 estimates (M = 18.4, SE = 1.9) and the TPA^{s} GT1M estimates (M = 16.4, SE = 1.8, t(22), $p < 0.01$, r = 0.5). Application of a correction factor to the TPA^{s} GT1M data meant that there was a non-significant difference between estimates from the TPA^{s} 7164 data (M = 18.4, SE = 1.9) and the TPA^{s} GT1M^{corr} data (M = 18.0, SE = 1.8, t(22), p = 0.68, r = 0.08). There was a non-significant difference between the TPA^{pu} 7164 estimate (M = 25.0, SE = 2.3) and the TPA^{pu} GT1M estimate (M = 23.9, SE = 2.3, t(22), p = 0.12, r = 0.3) and between the TPA^{pu} GT1M estimate and the TPA^{pu} GT1M^{corr} (M = 26.2, SE = 2.5, t(22), p = 0.13, r = 0.3).

The estimates from the 7164 and the GT1M were then compared with the CARS as the criterion method. Table 7.6 presents the mean difference and LOA (when applying the Sirard et al. (2005) cut-points) between the estimates, comparing the 7164 accelerometer and the GT1M accelerometer against the CARS criterion method.

Table 7.6: Mean difference (LOA) of minutes in sedentary behaviour and physical activity with the Sirard et al. (2005) cut-points.

Intensity	7164 *dm* (LOA)	GT1M *dm* (LOA)
Sed^{s} versus CARS	-4.0 (-20.3 to 12.4)	-5.9 (-20.1 to 8.3)
LPA^{s} versus CARS	7.5 (-5.3 to 20.3)	6.4 (-7.1 to 19.8)
MVPA^{s} versus CARS	-3.6 (-12.6 to 5.5)	-0.5 (-5.5 to 4.5)
TPA^{s} versus CARS	3.9 (-12.1 to 20.1)	5.9 (-8.3 to 20.1)

CARS: children's Activity Rating Scale; dm: mean difference; LOA: limits of agreement; LPA: light physical activity; MVPA: moderate-to-vigorous physical activity; s: Sirard et al (2005) cut-point; Sed: sedentary behaviour; TPA: total physical activity.

Wilcoxon tests were carried out with the level of significance set at p < 0.01 following the Bonferroni correction. Using the Sirard et al. (2005) cut-points the estimate from the GT1M accelerometer was significantly different from the CARS (p <0.001) except for the MVPA^{s} estimate (*Mdn* = 4.3) which was not significantly different from the CARS (*Mdn* = 3.5, z = -0.56, p = 0.58, r = -0.08). For the 7164 monitor there was a non-significant difference between the estimate of sedentary behaviour (Sed^{s}) (*Mdn* = 27.0) and the CARS (*Mdn* = 24.0, z = -2.05, p = 0.04, r = -0.30) and between the TPA^{s} estimate (*Mdn* = 16.3) and the CARS (*Mdn* = 20.0, z = -2.01, p = 0.04, r = 0.30). There were significant differences between the LPA^{s} estimate (*Mdn* = 9.0) and the CARS (*Mdn* = 15.8, z = -4.02, p = 0.00, r = -0.59) and MVPA^{s} estimate (*Mdn* = 6.5) and the CARS (*Mdn*= 3.5, z = -3.45, p = 0.001, r = -0.51).

The mean difference and LOA between the 7164, GT1M and the CARS with the Puyau et al. (2002) cut-points are presented in Table 7.7.

Table 7.7: Mean difference (LOA) minutes in sedentary behaviour and physical activity with the Puyau et al. (2002) cut-points.

Intensity	7164 *dm* (LOA)	GT1M *dm* (LOA)
Sed^{pu} versus CARS	2.7 (-12.9 to 18.3)	1.5 (-14.3 to 17.4)
LPA^{pu} versus CARS	0.7 (-10.7 to 12.1)	-1.5 (-15.1 to 12.1)
MVPA^{pu} versus CARS	-3.4 (-13.4 to 6.6)	-0.1 (-6.6 to 6.4)
TPA^{pu} versus CARS	-2.7 (-18.3 to 12.9)	-1.6 (-17.4 to 14.2)

CARS: Children's Activity Rating Scale; dm: mean difference; LOA: limits of agreement; LPA: light physical activity; MVPA: moderate-to-vigorous physical activity; pu: Puyau et al. (2002) cut-points applied; Sed: sedentary behaviour; TPA: total physical activity.

The results suggest that the smallest mean differences between estimates of time spent at different intensities of physical activity and the CARS and between estimates of time spent in sedentary behaviour and the CARS, were for the GT1M using the Puyau et al. (2002) cut-points. The results revealed that there were no significant differences between the GT1M estimates of minutes in physical activity and sedentary behaviour (Sed^{pu}, LPA^{pu}, MVPA^{pu}, TPA^{pu}) and the CARS (p > 0.001), with small effect sizes demonstrated (r = -0.007 to -0.19). There was a non-significant difference for the 7164 estimates for sedentary behaviour (Sed^{pu}) (*Mdn* = 19.5, z = -1.67, p = 0.09, r = 0.25) and for TPA^{pu} (*Mdn* = 23.5, z = -2.21, p = 0.03, r = 0.30) against the CARS (*Mdn* = 24.0 and 20.0 min, for time spent in sedentary behaviour and TPA respectively). The small and non-significant overall mean differences between the CARS estimates and the Puyau et al. (2002) estimates, could suggests that group level estimates of physical activity and sedentary behaviour may be sufficiently accurate. However, there were large limits of agreement for all the estimates made. This has important implications, as it suggests a limitation of the cut-points to accurately measure physical activity and sedentary behaviour at an individual level.

In summary, there was strong correlation between the monitors during mechanical calibration and both accelerometers displayed a similar pattern of counts when subjected to increasing accelerations. However, there were wide limits of agreement and a positive bias, with the mean difference not being close to zero. During free-play the cpm for the GT1M were significantly lower than for the 7164 accelerometer, with wide limits of agreement and a positive bias, for example the average cpm for the 7164 model were higher than the GT1M accelerometers. The time spent at different intensities was not comparable between the accelerometer models and this did not change for sedentary behaviour, LPA, or MVPA when a correction factor was applied to the GT1M data, or when the GT1M thresholds were reduced by 10%. Application of a correction factor to the GT1M data did however make the output for TPA comparable between the 7164 and the GT1M accelerometers.

Comparison of the estimates against the criterion measure supports the finding in an earlier chapter of this thesis that the Puyau et al. (2002) cut-points for the GT1M are most accurate at a group level when compared against the CARS criterion method.

7.4 DISCUSSION

The results suggest that both models had high intra-unit reliability, with an ICC value of > 0.8 (Matthews 2005). The CV and the *SEM* of the 7164 were larger than the GT1M, suggesting it had lower intra-instrument reliability, which is similar to the findings of earlier mechanical calibration studies (Rothney et al. 2008). While there was a significant correlation between models, there was a mean difference of 3.2 counts/s. This could result in a difference of 191 cpm which may not be biologically meaningful when considering cut-points for MVPA, such as ≥ 2296 cpm (Evenson et al. 2008), one of the thresholds recommended for use in children and adolescents (Trost et al. 2011).

The results of this study also suggest that there was a positive correlation between the cpm from the GT1M and the 7164 accelerometer during free-play. However, the bias of 632 cpm for the uncorrected GT1M data and 495 cpm for the corrected GT1M data were not close to zero and there were large limits of agreement. The scatter plot of the data revealed that the data for the GT1M were consistently lower than 7164 accelerometry data. It was also apparent from the Bland and Altman plot that the mean differences were accentuated at higher output (e.g. at higher intensity of physical activity). This possibly suggests that at higher intensities of physical activity the differences between the output of the different accelerometry models is accentuated.

The bar chart of the median cpm for each child illustrates the problem in comparing the output of the two generations of monitors as there were consistently higher median cpm for all children with the 7164 accelerometers. For two participants there was a difference of > 2500 cpm between the median values for the 7164 and GT1M accelerometers (participant 5 and 17). This highlights the concern regarding the comparability of different generations of accelerometers and raises questions regarding the ability to compare the findings of studies which have used different generations of Actigraph accelerometers.

In examining the time spent at different intensities, the GT1M recorded significantly less time in MVPA than the 7164 accelerometer. Applying a correction factor or reducing the threshold by 10% for the GT1M accelerometer did not change the difference for MVPA activity. This would result in a 52% difference in time spent in MVPA between the GT1M and the 7164 when the Puyau et al. (2002) cut-points are applied and would be of concern if this pattern continued over a 10 - 12 hour day of activity. As this study was limited to 1 hour of time-tabled unstructured free-play, where children could run and play, it might not necessarily reflect 'typical' activity over a day for a pre-school child. It was not possible in this study to determine whether these differences would continue over an extended period of data collection.

The comparability between models for estimates of time spent in sedentary behaviour and LPA depended on which cut-points were applied. There was a significant difference in the time spent in sedentary behaviour between the GT1M estimate and the Sirard et al. (2005) cut-points, with the GT1M recording significantly more time in sedentary behaviour. However, there was a non-significant difference for time when the Puyau et al. (2002) cut-points were applied. The lower threshold for the Puyau et al. (2002) cut-points (< 200 counts/15 s) for sedentary behaviour could result in less time being classified as sedentary for the GT1M and reducing the Sirard et al. (2005) cut-point by 10% may explain why this difference then became non-significant.

Using the Sirard et al. (2005) cut-points there was a non significant difference for estimates of TPA with the GT1M once the correction factor had been applied. Prior to application of the correction factor the GT1M was underestimating total physical activity in comparison with the 7164.

In the previous chapter, the difference in estimates of time spent in different intensities was shown to vary depending on which cut-point is applied. To investigate whether the application of a correction factor would be of value to allow cross-comparison between accelerometry models for TPA, the only other available cut-point was investigated, which allows calculation of total physical activity and which has been developed for young children, by Van Cauwenberghe et al. (2011). While other cut-points have been validated for pre-school children, they only provide thresholds for MVPA (Jimmy et al. 2012; Pate et al. 2006), and therefore it is not possible to estimate TPA. The results are presented in Table 7.8.

Table 7.8: Estimate of time spent in total physical activity using the Van Cauwenberghe et al. (2011) cut-points.

	Time min (SD)		
	7164	**GT1M**	**GT1M^{corr}**
Van Cauwenberghe et al. (2011)	17.8 (9.2)*	15.7 (8.3)	17.2 (9.1)

*Values in bold not significantly different from 7164; * values significantly different from 7164 estimate.*

This suggests that the application of a correction factor to the GT1M data may improve comparability between the output from the 7164 accelerometer and the GT1M accelerometer for TPA in pre-school children.

In conclusion, this study supports the findings of the previous study in this thesis (Chapter 6) that the Puyau et al. (2002) cut-points for the GT1M data have the ‘best’ agreement with the CARS criterion measure. However, if researchers are interested in the convergent validity of the GT1M with the 7164 accelerometer, applying a correction factor to the GT1M data (7164 = GT1M/0.91) may be appropriate for estimates of total physical activity. The correction factor may partially resolve the issue of the output from the GT1M having lower cpm than the 7164 model, which may result in more time being classified as ‘sedentary’ and less time in LPA, MVPA or TPA. However, as this study was limited to data collected from 23 participants over 1 hour of direct observation, further investigation is warranted. In particular, it would be important to examine whether the correction factor improves convergent validity, for free-living physical activity data collected from a larger population over an extended period of time.

CHAPTER 8 : AN INVESTIGATION INTO THE MINIMUM WEAR TIME TO RELIABLY ESTIMATE HABITUAL PHYSICAL ACTIVITY IN PRE-SCHOOL CHILDREN

8.1 INTRODUCTION

Despite accelerometry being used extensively as an objective means of measuring physical activity in young children (Pate et al. 2010) there are several outstanding questions on approaches to reduce accelerometry data, whereby accelerometry output is transformed into a meaningful format (Cliff et al. 2009b). As part of the data reduction process, researchers have to make a number of methodological decisions for which there are different approaches and recommendations and this complicates the process (Oliver et al. 2012). Currently there are inconsistencies between studies in the methods they adopt for data reduction and data processing, for example, how missing data are handled, and whether this data are imputed, or not, varies between studies. Similar to the problems already discussed with the application of different cut-points, there is a need for standardisation in the data reduction processes between studies to allow for meaningful interpretation and synthesis of the findings of different studies (Mâsse et al. 2005).

8.1.1 Valid day and hours

Two of the key decisions in data reduction are to determine how many hours of data constitute a 'valid' day and how many days will provide a reliable estimate of habitual physical activity and time spent in sedentary behaviour (Ojiambo et al. 2011). This has an impact on whether to subsequently exclude participant data with insufficient days or hours of data. In excluding incomplete participant data, there is a risk of potential inaccuracies, which can result from differences between days that are excluded and days kept in the analysis. Therefore, the decision whether to exclude incomplete participant data or include data for part of a day can produce biased estimates of physical activity and result in different conclusions being reached (Alhassan et al. 2008).

To date, there is limited evidence of the number of days and hours within days required to provide a reliable estimate of physical activity in the pre-school population (Hinkley et al. 2012b). There have been two studies with pre-school children which have explored how many days and hours of data were necessary to provide a reliable estimate of habitual physical activity. Penpraze et al. (2006) conducted reliability analysis on 7 days of accelerometry data collected from 76 pre-school children (mean (SD) age: 5.6 (0.4) y). The authors argued that while 10 hours a day over 7 days maximised the reliability of estimates of total physical activity (r = 0.80), 3 days of data had sufficient reliability ($r \geq 0.60$). They also suggested that the increase in reliability coefficients between 3- and 10-hours of data collection were small ($r \sim = 0.02$), when 1 to 7 days of monitoring were used. In a more recent study by Hinkley et al. (2012b), reliability analysis was conducted on data collected from 1004 pre-school children (aged 3 to 5 years), over an 8-day period. The authors concluded that the number of days of data required to achieve a desired Intraclass Correlation Coefficient (ICC) estimate of 0.7, increases as the number of hours of data per day decreases. For example, in their study, 10 hours of data per day required 2.8 days of data collection, while collection of 7 hours of data per day meant that 3.4 days were required to achieve the same ICC estimate.

Hinkley et al. (2012b) caution that when making decisions on the number of hours or days of data used there needs to be a trade off between using stringent inclusion criteria for accelerometry data and ensuring an adequate sample size. For example, increasing the desired ICC value would require an increase in the number of days and hours of data collection to achieve the target. As a consequence fewer participants will have 'complete' data resulting in their exclusion from analysis and thus reducing the overall study sample size.

The commonly adopted approach to determine how many hours constitute a valid day and how many days reflect an individual's usual or 'habitual' physical activity, is to first ascertain the intra-individual variability of activity between days (Baranowski and de Moor 2000). ICC calculations can be used to estimate the consistency of activity across days (Baranowski and de Moor 2000). If an individual replicates the same activity pattern every day then 1 day of data would be sufficient. However, as variability increases across days then an increasing number of days are necessary to reflect habitual activity. Similarly, the number of hours to reflect a 'typical' day also contributes to the variability in data between days.

One important consideration in determining habitual activity is whether the inclusion of a weekend day increases variability, as a result of the activity patterns being different to weekdays. This consideration impacts on whether it is necessary only to include participant data that includes a weekend day, or conversely exclude participant data that does not include a weekend day. Both the studies by Hinkley et al. (2012b) and Penpraze et al. (2006) investigated the implications of including a weekend day in data analysis. Penpraze et al. (2006) reported significant differences in the mean cpm between weekdays and weekend days. However, the inclusion of a weekend day had minimal impact on the reliability of the data and the authors suggest that, in their sample, the inclusion of a weekend day in analysis is not necessary. In contrast, Hinkley et al. (2012b) also reported that total physical activity differed significantly between weekdays and weekend days. The authors recommend that future studies ensure both weekend and weekdays are included to accurately represent pre-school children's physical activity across an entire week.

8.1.2 Non-wear time

In addition to identifying the number of days and hours of data required, data reduction also requires decisions to be made on how periods during the day are identified when the accelerometer is not worn (non-wear time) and how these missing data are then accounted for in the analysis. Over several days of data collection, missing data may be due to the removal of accelerometers e.g. during water-related activities, for nap-times or for non-compliance with protocols for wearing the accelerometer. Accelerometers are usually worn during waking hours and when removed they will record a zero count over an epoch. Extended periods of consecutive zeros in habitual physical activity data may be flagged up as behaviourally unlikely and attributed to the accelerometer being removed. However, it is argued that sedentary behaviour such as sitting may also record zero counts (Evenson and Terry Jr 2009) and without subjective information it is not possible to determine whether the accelerometer was removed or not. While some researchers recommend participants keep a log when an accelerometer is removed (Esliger et al. 2005; Trost et al. 2005), this introduces an element of self-report which can be problematic particularly in research with young children (Sallis and Saelens 2000). Asking participants to document details of wear time also adds an additional participant burden which may hinder compliance within a study (Evenson and Terry Jr 2009).

An alternative to keeping a log is to deduce non-wear time from scanning the accelerometry data to identify extended periods of consecutive zeros. Several studies have considered the issue of non-wear time in adult populations (Choi et al. 2011; Evenson and Terry Jr 2009; Winkler et al. 2009) and have identified thresholds, described as a number of minutes of consecutive zero counts in the accelerometry data, in order to identify periods of non-wear time. Currently, the recommended criteria for non-wear time in adult populations is a threshold between 60 - 90 minutes with the inclusion of up to 2 minutes of non zero counts either side (Oliver et al. 2012). The two minutes is to allow for any momentary accidental movement of accelerometers once removed which may interrupt a period of consecutive zeros but which would not constitute true wear time. Studies have gone on to apply different thresholds as a means of detecting non-wear time and this has led to the exclusion of 10 minutes of consecutive zeros (Eiberg et al. 2005; Hinkley et al. 2012b; Mattocks et al. 2008), 20 minutes of consecutive zeros (Alhassan et al. 2008; Esliger et al. 2005; Ojiambo et al. 2011; Treuth et al. 2004a), 60 minutes of zeros (Matthews et al. 2008) and up to 180 minutes of consecutive zeros (van Coevering et al. 2005) in different studies with children and adolescents.

Although the accuracy of different thresholds for identifying episodes of non-wear time has been explored in adult populations against criterion methods (King et al. 2008; Oliver et al. 2012), there are currently only two studies with school-aged children (Alhassan et al. 2008; Esliger et al. 2005) which have investigated appropriate thresholds of non-wear time. Esliger et al. (2005) reported on data from a sample of 115 children (8 to 13 y) in which they explored the length of longest time period of motionless data (periods of consecutive zeros). The authors found that 90% of the periods of motionless data were less than 10 minutes and 76% of the participants had longer periods. The mean (95% CI) of the longest periods of consecutive zeros was 17.5 (1.5) min. Alhassan et al. (2008) also suggest from results of pilot data that a 20 minute period of consecutive zeros was the maximum number of minutes of zeros associated with inactivity. A limitation with both studies is that neither study has validated the proposed thresholds against a criterion method and there are no methodological details of the pilot study undertaken by Alhassan et al. (2008). Nevertheless, Esliger et al. (2005) and Alhassan et al. (2008) both support the use of a threshold of $\geq$ 20 minutes of consecutive zeros to identify periods of non-wear time in children. The authors suggest that any longer periods of consecutive zeros would be behaviourally improbable, as it is unlikely to reflect sedentary activity but rather suggest that the accelerometer had been removed.

Once non-wear time has been identified there are different approaches to handling the data: one is to ignore the missing data and to keep this in the analysis (Esliger et al. 2005), a second approach is to exclude the data from the final analysis, and a final option is to impute the missing data (Cliff et al. 2009b). One of the implications of ignoring non-wear time and including extended periods of zero counts in the analysis is that it could dilute the average cpm over the data collection period (Esliger et al. 2005). This may result in an underestimation of true levels of physical activity as a consequence. In contrast, excluding periods of non-wear time can mean that participants potentially have incomplete days of data. This could impact on whether a participant has sufficient days or hours of data to be included in the analysis. As a consequence, sample size could be affected and this could increase the risk of sampling-bias having an impact on the results (Mâsse et al. 2005).

Several approaches have been proposed to impute missing data, where observed data values are used to assist in prediction of missing values (Catellier et al. 2005; Esliger et al. 2005; Mâsse et al. 2005). However, it is argued that many of these approaches are not feasible for research teams to use on accelerometry data or require further empirical studies to determine their credibility (Cliff et al. 2009b). One approach by Catellier et al. (2005) is to define a 'standard day', where there are non-missing counts over at least 80% of a 'standard measurement day'. A 'standard measurement day' is defined as the length of time when at least 70% of the sample of participants are wearing their monitor (Catellier et al. 2005). The value of this approach has not been explored with pre-school children; in addition, the implication of applying different thresholds of non-wear time has not been explored with this population.

8.1.3 Summary

In summary, determining the number of hours and days of data to reliably reflect habitual physical activity needs further investigation in the pre-school population. This may be sample-specific and therefore each study needs to determine what is appropriate for its own unique sample. Whether having criteria for the inclusion of a weekend day in the analysis needs further examination as there are conflicting opinions on this. Finally, despite the different approaches to dealing with non-wear time, no consistent or standardised approach has been adopted across studies. The implications of applying different criteria for non-wear time has not been explored with pre-school children. It is important to determine whether the application of different non-wear time criteria influences estimates of physical activity or sedentary behaviour and to investigate how these factors will impact on sample sizes.

The aims of this study were:

- To determine the recommended wear time required to provide reliable estimates of habitual physical activity and sedentary behaviour of pre-school children.

- To investigate whether the inclusion of a weekend day is necessary for reliable estimates of habitual physical activity and sedentary behaviour of pre-school children.
- To examine the influence of applying non-wear time criteria to estimates of physical activity and sedentary behaviour of pre-school children with 7 days of free-living accelerometry data.

8.2 METHOD

8.2.1 Data collection

This study involved a secondary data analysis of 121 accelerometry data files collected from children aged 42 - 48 months. The analysis is based on baseline accelerometry data collected at the start of a longitudinal intervention study with pre-school children. The baseline data were collected between November 2009 and July 2010. Recruitment was undertaken from two demographically similar geographical areas within Scotland. One area was used for recruitment of an intervention group and the other for recruitment of a control group. To determine sample size at the outset it was estimated that 110 participants in total (55 per group) were necessary to detect changes of 100 cpm per day powered at 80%. Children were recruited from nurseries in the two geographical areas. To ensure recruitment of participants from a range of socioeconomic backgrounds, postcodes for the Nurseries were used to allocate a Scottish Index of Deprivation (SIMD) quintile (with 1 being most deprived and 5 least deprived quintile (Scottish Government 2010)). In addition, children were selected that had postcodes which matched closely with the nursery to try to ensure there was sampling across the quintiles.

Data in this analysis had been collected using the GT1M accelerometers and the GT3X accelerometers set to uniaxial mode, which were distributed to parents of pre-school children attending nurseries in the identified areas. Parents were asked to attach the accelerometers, which were on an elasticated belt, around their children’s waists during waking hours for 7 consecutive days. The accelerometers were set to collect data in 15-s epochs. Parents were asked to remove the accelerometers for any water based activities (bathing, swimming). Data were transferred to Excel and processed using the MAHUffe software (MRC epidemiology unit 2013).

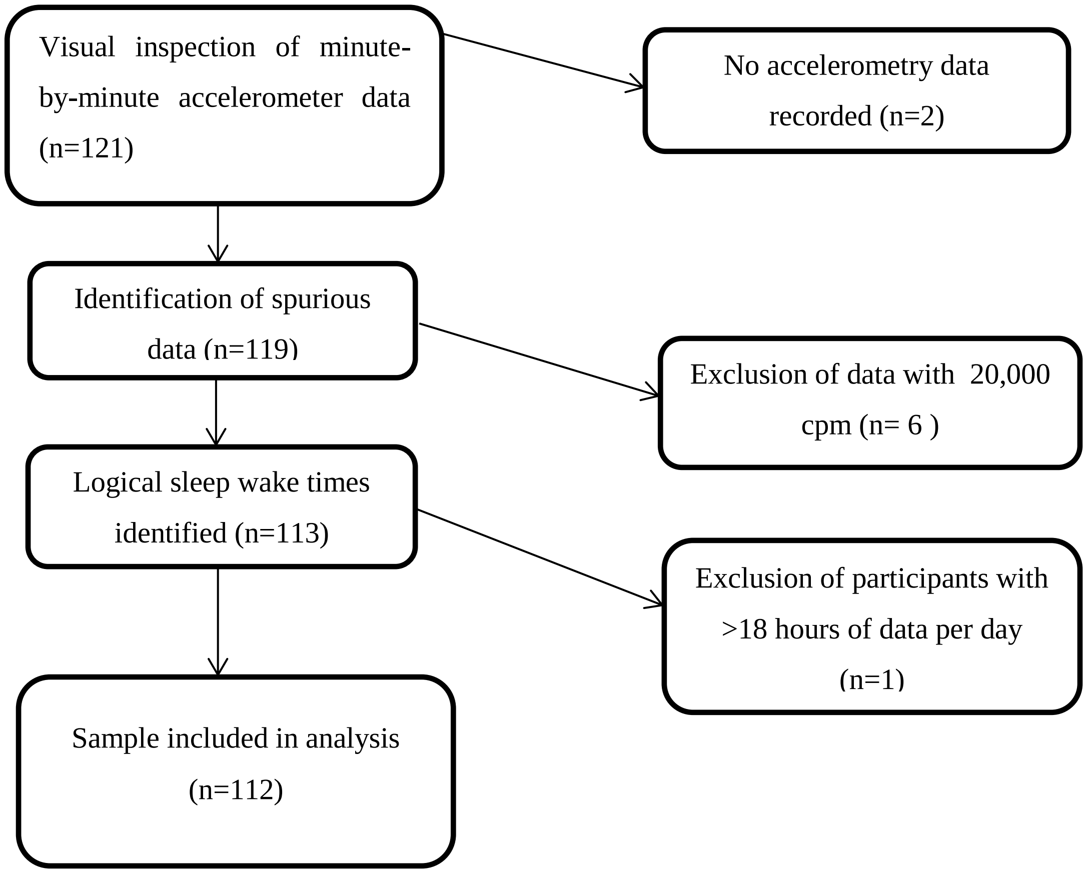
8.2.2 Data cleaning

The initial step in data reduction was to clean the 121 accelerometry data files. Each file represented a participant's accelerometry data collected over 7 days. The first consideration was to identify spurious data (data with very high data counts). Minute-by-minute visual inspection of accelerometry data output files was undertaken and spurious data were considered to be when the cpm was > 20,000 for a 1-minute epoch. The MAHUffe programme presents an individual's accelerometry data in a graph as well as allowing for minute-by-minute inspection of the data. APPENDIX VIII: present examples of where the cpm data were < 20,000 cpm (Appendix Figure VIII.i) and > 20,000 cpm (Appendix Figure VIII.ii). It can be seen that in Appendix Figure VIII.ii that the cpm were consistently > 20,000 cpm (reaching as high as 200,000 cpm) which possibly suggests an error with the monitor (see paragraph below) and this participant's data was excluded.

The upper limits of biological plausibility for accelerometry data have been suggested as when cpm are ≥ 15,000 (based on unpublished observation data of 94 youth 8 to 13 y) (Esliger et al. 2005). Using a linear regression relationship between treadmill speed and Actical accelerometry counts Colley et al. (2010) extrapolated data to identify a threshold of > 20,000 cpm as being spurious data, which would not be biologically plausible. In the current study visual inspection of the days with > 20,000 cpm revealed a particular error pattern with data presenting with mean daily cpm of >50,000 cpm. It is possible that this may relate to an error code indicating monitor malfunction which has been reported in earlier studies (Alhassan et al. 2008). Alhassan et al. (2008) reported an error code of 32,767 to represent voltage signal saturation with the 7164 Actigraph monitor. In the current study the error pattern of mean daily cpm being > 50,000 cpm was apparent in data collected from six participants, and the data from these participants were excluded from the analysis.

Visual inspection also revealed two participants who had no accelerometry data and one participant who had worn the accelerometer for > 18 hour/day. This participant's data was excluded as it did not correspond to logical sleep/wake time (Alhassan et al. 2008). This left a total of 112 participants. Figure 8.1 presents the stages of data cleaning.

Figure 8.1: Stages of data cleaning.



8.2.3 Data analysis

1) Calculation of a 'standard day'

To allow comparison of weekdays with weekend days, the days of the week were identified and separated for analysis. The start and finish times for each participant (n = 112) were identified and frequency plots for all participants was undertaken. Table 8.1 presents the characteristics of the sample, with 78% classified as 'healthy' weight and 22% classified overweight or obese i.e. BMI at or above 85% centile relative to the UK population reference data (Cole 2002)). Data with ≥ 20 minutes of consecutive zeros were identified as non-wear time and excluded from analysis as recommended by Esliger et al. (2005). A 'standard measurement day' was identified as the length of time when at least 70% of the sample wore the accelerometer for 80% of the time (Catellier et al. 2005).

Table 8.1: Characteristics of sample (Chapter 8).

	Mean (SD)		
	All	**Male**	**Females**
No. participants	112	60	52
Age (years)	3.7 (0.7)	3.7 (0.7)	3.7 (0.7)
Height (cm)	101.3 (3.9)	101.6 (3.7)	101.0 (4.1)
Body weight (kg)	17.0 (2.3)	17.5 (2.3)	16.5 (2.1)
Body mass index (kg/m^2)	16.6 (1.6)	17.0 (1.7)	16.2 (1.3)

2) Comparison of weekend and weekdays

Data from all 112 participants (60 males, 52 females, mean (SD) age: 3.7 (0.2) y) were included in the analysis of weekend and weekdays. Data with ≥ 20 minutes of consecutive zeros were identified as non-wear time and excluded from analysis (Esliger et al. 2005).

Using all participant data the mean cpm for participant's weekdays of data and the mean cpm for their weekend days were extracted to allow comparison between weekend and weekdays. Data were evaluated for normality using the Kolmogorov-Smirnov as the sample size was more than 50. Results of the normality tests are provided in Appendix VIII (Appendix Table VIII.i). As the data were not normally distributed, descriptive statistics of the medians and inter-quartile ranges (IQR) were calculated. Box plots giving the median and IQR were used to present the data graphically. The Wilcoxon signed rank test, the non-parametric equivalent of the paired *t*-test was used to compare differences between cpm for weekend and weekdays. In addition, a comparison was made between the cpm for males and female at weekend and weekdays using an Mann-Whitney U test, the non-parametric equivalent of the independent *t*-test. The level of significance was set at $p < 0.05$. Effect sizes were calculated to provide an objective measure of the importance of an effect using the Pearson's correlation coefficient r (Field 2012). The effect size was calculated from z-scores using Equation 8.1.

Equation 8.1: Calculation of effect size

$$\frac{z - score}{\sqrt{number\ of\ comparisons}}$$

(Field 2012)

The *r* effect sizes range from 0, suggesting no effect to 1 suggesting a perfect effect. Table 8.2 presents the effect size criteria proposed by Cohen (1988; 1992) for small and large effect sizes which are widely accepted (Field 2012).

Table 8.2: Interpretation of effect sizes.

r	***Categorisation of effect***	***Explanation of effect***
r = 0.10	Small effect	explains 1 % of the variance
r = 0.30	Medium effect	explains 9 % of the variance
r = 0.50	Large effect	explains 25 % of the variance

(Adapted from Field 2012)

Using the MAHUffe programme (MRC epidemiology unit 2013) the data were processed by applying cut-points for sedentary (< 200 counts/15 s) and TPA (using the threshold of ≥ 200 counts/15 s) as defined by Puyau et al. (2002). These cut-points have been validated during free-play of pre-school children against direct observation (Hislop et al. 2012a). A comparison was made between the percentage time spent in TPA between weekdays and weekend days. The mean percentage of time spent in TPA was calculated for participants with 4 or more weekdays and for those with 1 or more weekend days. The criteria of 4 days was selected as this is frequently used in studies of pre-school children (Alhassan et al. 2008; Cardon and Bourdeaudhuij 2008; Penpraze et al. 2006). The estimates of mean TPA was calculated for each hourly increment from 3 hour/day to 10 hour/day. The percentage of time spent in TPA was determined for each day of data by dividing the number of minutes spent in TPA by the total wear time for that day and multiplying by 100 (Hinkley et al. 2012b). This approach takes into consideration any possible differences in wear time within and between participants (Hinkley et al. 2012b). The Kolmogorov-Smirnov statistic was used to assess the normality of the data as the sample size n > 50. The results of the normality test suggested that not all data were normally distributed (Appendix VIII, Appendix Table VIII.i). The non-parametric test, Friedman's ANOVA was therefore used to assess the difference between weekend and weekday total physical activity.

3) Reliability analysis

Data from all 112 participants (60 males, 52 females, mean (SD) age: 3.7 (0.2) y) were also included in the reliability analysis and ≥ 20 minutes of consecutive zeros was identified as non-wear time and excluded from analysis (Esliger et al. 2005).

Participants who had at least 4 days of data, described as being 'any' 4 days of data (which could include a weekend day) were identified for the reliability analysis. In addition, participants with 4 days which included at least 1 weekend day were also identified. Days with data ranging from 3 or more hours per day of data to 10 or more hours of data per day were entered into the analysis.

A repeated measures ANOVA was conducted to investigate any sources of systematic bias between days of data collected, including weekend days, and to detect any non random change between days within the data. This test is useful in detecting large systematic bias relative to the random error (Atkinson and Nevill 1998).

The dependent variable was the time spent in TPA and the independent variable was the day of testing. The *F* value represents the ratio of the systematic variance to the unsystematic variance (Field 2012).

To calculate the consistency of activity across days an ICC was calculated. An ICC (2,1) was selected where days was identified as the random effect, as the aim was to generalise the results from random days in the analysis to the other days within the sample population (Batterham and George 2000).

An ICC (2,1) was calculated using Equation 8.2 for each of the hours, where σ_b^2 is the '*between participant*' variance and σ_w^2 is the '*within participant*' variance.

Equation 8.2: Calculation of ICCs:

$$\sigma_b^2 / \sigma_b^2 + \sigma_w^2$$

(Trost et al. 2005)

Hopkins et al. (2000) states that at least 50 participants performing three or more trials provides adequate precision in estimating typical error.

The Spearman-Brown prophecy formula (S-B prophecy formula) was used (Stanley and Angoff 1971) to estimate the number of hours and days of data to achieve an *acceptable* level of reliability. Baranowski and de Moor (2000) argue that an ICC of ~>0.8 is an *acceptable* level of reliability. ICC single day measures were calculated for participants with any 4 days of data and repeated for participants with 4 days including 1 weekend day. ICC values were calculated for 3 or more hours, up to 10 hours per day of data and with 1 or more and up to 7 days of data. For each participant 4 days of data were then randomly selected for analysis of 'any' days. This was repeated with 4 days of randomly selected days which included at least 1 weekend day.

Using these values the S-B prophecy formula was used to estimate the hours and days required to achieve the target ICC values of 0.7, 0.8 and 0.9 using Equation 8.3.

Equation 8.3: Spearman-Brown prophecy formula

$$N=[ICC_{target}/(1-ICC_{target})]\ [(1-ICC_{s})/ICC_{s}]$$

(Stanley and Angoff 1971)

Where, N is the number of days and ICC_{target} is the ICC value desired, and ICC_{s} is the single day measure of reliability. Once the number of days was calculated the number of children achieving the criteria was calculated. The number of days required was rounded up when a decimal value of number of days was calculated such that 3.2 and 3.7 days would both be rounded up to 4 days, to ensure the minimum reliability criteria were met.

To compare the effect of a weekend day on reliability, reliability coefficients and 95% confidence intervals were also calculated (Equation 8.4) similar to the study by Penpraze et al. (2006) for 4 days of data for weekdays only and for 4 days including 1 weekend day:

Equation 8.4: Calculation of reliability coefficient

$$R = \frac{\sigma_b^2}{\left[\sigma_b^2 + \left(\frac{\sigma_w^2}{n}\right)\right] * 100}$$

(Adapted from Penpraze et al. 2006)

σ_b^2 is the *between participant* variance and σ_w^2 is the *within participant* variance and n equals the number of days.

4) Comparison of different criteria for non-wear time

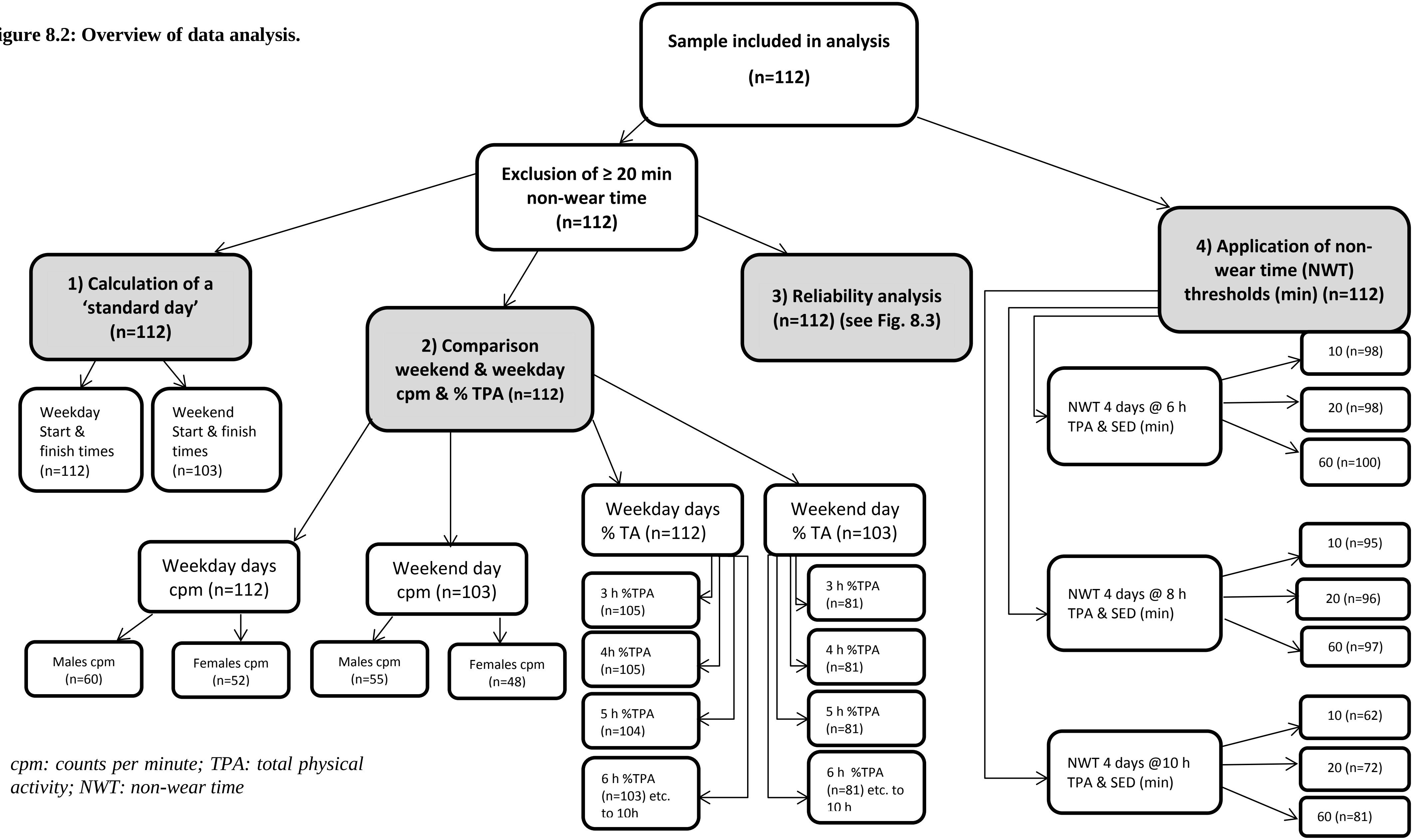
Using all data from the total sample (n = 112), the influence of different threshold criteria of 10, 20 and 60 minutes of non-wear time on participant involvement for 6, 8 and 10 hours of data collection was calculated.

The data concerning mean minutes of total physical activity and sedentary time for participants with at least 4 days of data were also analysed using the different criteria of 10, 20 and 60 minutes of non-wear time at the thresholds of 6, 8 and 10 hours of data collection. The results of the normality tests are presented in Appendix VIII (Appendix Table VIII.i). The data concerning mean of TPA data were normally distributed ($p > 0.05$) and therefore the mean and SD were presented graphically using bar charts with error bars to indicate one SD. The data concerning sedentary time were not normally distributed ($p < 0.05$) and were therefore presented as box plots indicating the median and IQR.

As the same participant data was re-processed using different non-wear time criteria, this meant that the data were not truly independent of each other e.g. the data were 'nested' within each other. This means that the data violated the assumptions necessary for undertaking inferential analysis and therefore it was not statistically sound to test for differences between the outcomes of the data sets. As a result descriptive statistics on the impact of non-wear time are presented. This still allowed for clear comparison of the impact that different non-wear criteria would have on estimates of time spent in TPA and sedentary behaviour.

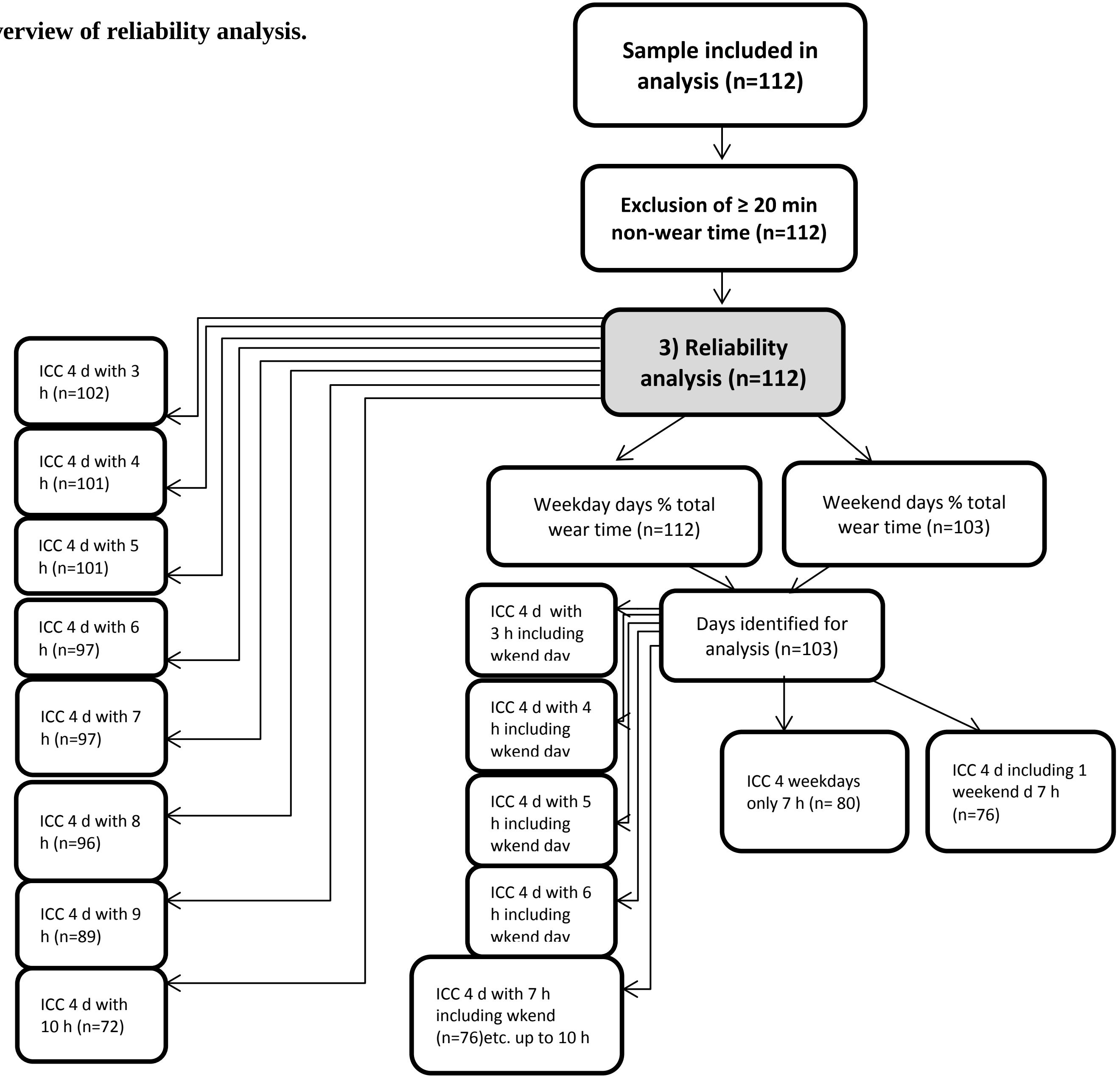
Figure 8.2 and Figure 8.3 present flow charts of the stages of analysis undertaken.

Figure 8.2: Overview of data analysis.



cpm: counts per minute; TPA: total physical activity; NWT: non-wear time

Figure 8.3: Overview of reliability analysis.



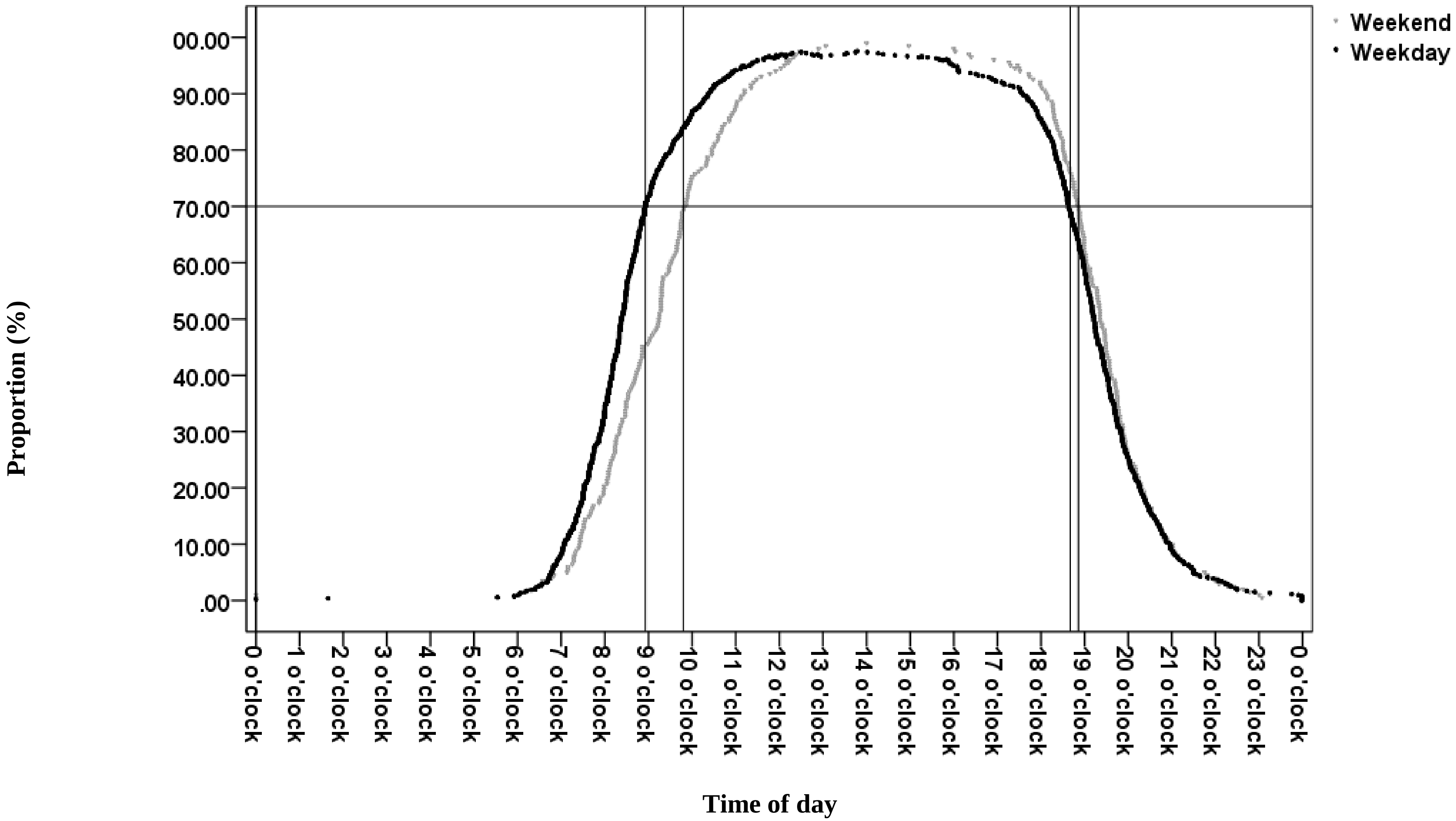
8.3 RESULTS

8.3.1 Standard measurement day

The results of the Kolmogorov-Smirnov statistic revealed that the data for registered wear time was not normally distributed (Appendix VIII, Appendix Table VIII.i) and therefore the median (IQR) were reported. The median (IQR) number of days of data collection was 7 (1) days with median (IQR) of 583.2 (85.1) minutes or 9.7 (1.4) hours of registered time. The median (IQR) hours of data collection for weekdays was 10.4 (1.5) hours and for weekend days this was 9.8 (2.2) hours.

Figure 8.4 presents a frequency plot of the start and finish times of all participants (n = 112), and shows the cumulative proportion of children identified as wearing their accelerometers at different times of the day. Reference lines on this figure give an indication when 70% of the sample wore accelerometers for 80% of the day, defined as a 'standard measurement day' (Catellier et al. 2005). For weekend days, the median (IQR) start time was 09:14 (08:11 to 09:59) and finish time was 19:21 (18:43 to 20:01). For weekdays the median start time was 08:23 (07:43 to 09:07) and finish time was 19:13 (18:31 to 20:01).

Figure 8.4: Standard measurement day.

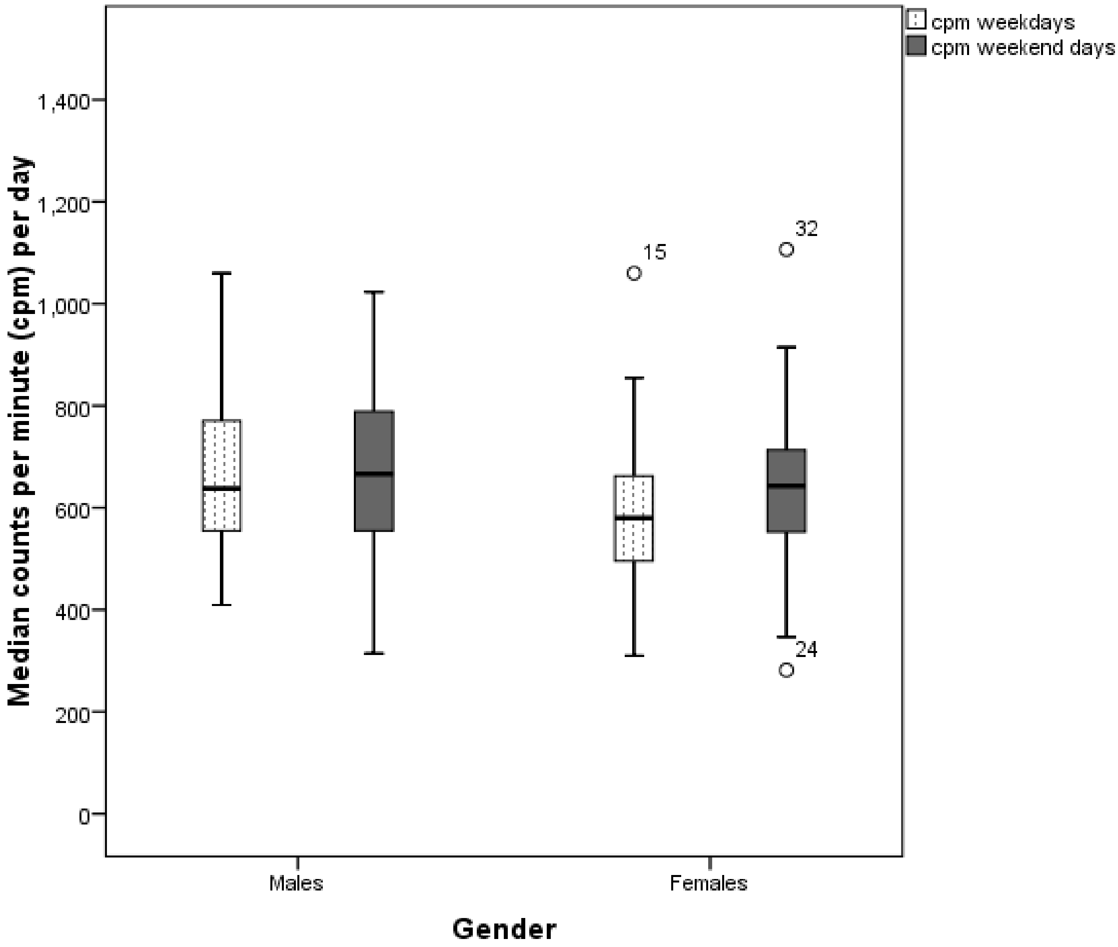


Using the plots from the graph a ‘standard weekday’ was from 08:55 to 18:40 and a standard weekend day was identified as 09:48 to 18:51. The results suggest that for 70% of population the median wear time [Standard measurement day = (finish time – start time)*0.8] for weekend days was 7 hours 14 min (434 min) and 7 hours 48 min (468 min) for weekdays. It is interesting to note that the frequency plot pattern of wear time is similar for weekend and weekdays. This is in contrast to the pattern seen in adolescent girls reported in the study by Catellier (2005), where wear time during weekdays had a markedly different pattern to wear time during a weekend day, whereby start times commenced later and finished later.

8.3.2 Comparison of weekend and weekdays

Data concerning the cpm for weekend and weekdays were not normally distributed and therefore the median (IQR) values are reported. The median (IQR) cpm for weekend days was 653.4 (220.3) cpm and for weekdays it was 607.4 (180.9) cpm. Results of the Wilcoxon Signed Rank test revealed that there was no significant difference between the weekend and weekday cpm (z = 1.66, p = 0.097 r = 0.11). Using the Mann-Whitney U test, there was a significant difference with a medium effect size between cpm between males (*Mdn* = 643.0 (213.5) cpm) and females (*Mdn* = 571.8 (157.1) cpm) for weekdays (z = -3.07, p < 0.05, r = -0.30), but no difference between males (*Mdn*= 665.8 (239.8) cpm) and females (*Mdn* = 643.1 (172.1) cpm) for weekend days (z = -1.42, p = 0.16, r = -0.14). There was also a non-significant difference for males between weekdays and weekend days (z = -3.44, p = 0.73, r = -0.21) but a significant difference for females (z = -2.18, p = 0.03, r = -0.35) between weekdays and weekend days. Figure 8.5 presents the median and inter-quartile range for the cpm for males and females comparing weekdays and weekend days.

Figure 8.5: Box plot of the median (IQR) of cpm for males and females comparing weekend with weekdays.



The percentage of time spent in TPA was calculated for all participants (n = 112) who had at least 4 days of data. This was calculated in hourly increments starting with participants who had at least 3 hours of data, then for participants with at least 4 hours of data, 5 hours of data up to participants with at least 10 hours of data. The Kolmogorov-Smirnov statistic showed this data to not be normally distributed ($p < 0.05$) (Appendix VIII, Appendix Table VIII.i). Weekend days with 4, 5, 7, 8, 9 and 10 hours of data were normally distributed ($p > 0.05$). Table 8.3 presents the median percentage (IQR) of percentage time spent in TPA for weekend and weekdays.

Table 8.3: Comparison of median (IQR) percentage time spent in TPA for participants with at least 4 weekdays and for participants with at least 1 weekend day.

Min number of hour/day	n	Median (IQR) % time in TPA Weekdays only (at least 4 days)	n	Weekend days only (at least 1 weekend day)
10 h	78	25.9 (9.4)	59	26.3 (9.0)
9 h	98	25.3 (9.6)	71	26.5 (10.2)
8 h	99	25.2 (9.6)	78	25.8 (10.6)
7 h	101	25.1 (9.5)	80	26.3 (10.7)
6 h	103	25.1 (9.5)	81	26.0 (10.8)
5 h	104	25.1 (9.5)	81	26.0 (10.6)
4 h	105	25.0 (9.6)	81	25.8 (11.3)
3 h	105	24.9 (9.7)	81	25.7 (11.8)

IQR: interquartile range; TPA: total physical activity.

Results of the Friedman's ANOVA suggests there was no significant difference between percentage time spent in TPA between weekend and weekdays (χ^2 (15) = 13.85 p > 0.05). The results suggest that the participants spent up to 1.2 % more time in TPA during weekend days in comparison to weekdays.

Test of systematic bias

Table 8.4 presents the results of the repeated measures ANOVA on the TPA recorded for each of the hours of data. There was no significant difference between tests suggesting no significant systematic bias. This is important as systematic bias can present a significant measurement error where factors influencing variation in physical activity across days could result in under- or over-estimation in activity levels across the sample (Ridley et al. 2009). However, it should be noted that one drawback of the ANOVA calculation is that the detection of systematic bias is affected by large random (residual) variation (Atkinson & Nevill, 1998).

Table 8.4: Results of repeated measures ANOVA for total physical activity.

Hours of data collection	Results of ANOVA
10 h	F (3,213) = 1.06, p = 0.368
9 h	F (3,264) = 0.43, p = 0.73
8 h	F (3,285) = 0.33, p = 0.8
7 h	F (3,288) = 1.34, p = 0.26
6 h	F (3,288) = 0.70, p = 0.55
5 h	F (3,300) = 0.44, p = 0.72
4 h	F (3,300) = 0.32, p = 0.81
3 h	F (3,303) = 0.46, p = 0.71

8.3.3 Reliability analysis

The results of the single day ICC measures and the S-B Prophecy formula for any 4 days of data are presented in Table 8.5. The results of 4 days including a weekend days are outlined in Table 8.6.

Table 8.5: Results of Spearman-Brown Prophecy formula based on any 4 days of data collection (which can include a weekend day).

Hours	No. of participants entered in analysis	Single day ICC	95% CI	No. of days required to achieve ICC values of:			No. (%) of children in sample (out of 112) meeting ICC values of:[a]	
				0.7	0.8	0.9	0.7	0.8
10 h	72	0.35	0.23,0.48	4.33	7.43	16.71	55 (49)	0
9 h	89	0.38	0.27, 0.50	3.81	6.53	14.68	89 (79)	32 (29)
8 h	96	0.44	0.34, 0.55	2.97	5.09	11.45	100 (89)	80 (71)
7 h	97	**0.44**	**0.34, 0.55**	**2.97**	5.09	11.45	**104 (93)**	85 (76)
6 h	97	0.43	0.33, 0.54	3.09	5.30	11.93	97 (87)	84 (75)
5 h	101	0.43	0.33, 0.53	3.09	5.30	11.93	101 (90)	88 (79)
4 h	101	0.43	0.33, 0.54	3.09	5.30	11.93	101 (90)	88 (79)
3 h	102	0.40	0.30, 0.51	3.50	6.00	13.5	102 (91)	92 (82)

[a]*None of the children met the criteria for reliabilities of 0.9*
The suggested 'optimum' ICC values for this sample are highlighted. Using the single day ICC value of 0.44, the S-B prophecy formula suggests that for 7 hours of data collection, for 3 days would achieve the desired ICC value of 0.7. This would include 93% of the original sample.

Table 8.6: Results of Spearman-Brown Prophecy formula based on 4 days including 1 weekend day.

Hours	No. of participants entered in analysis	Single day ICC	95% CI	No. of days required to achieve ICC values of			No. (%) of children in sample (out of 112) meeting ICC values of:[a]	
				0.7	0.8	0.9	0.7	0.8
10 h	54	0.27	0.12, 0.41	6.64	11.38	25.61	13 (12)	0
9 h	68	0.37	0.25, 0.51	3.97	6.81	15.32	68 (61)	28 (25)
8 h	74	0.35	0.23, 0.48	4.33	7.43	16.71	73 (65)	3 (3)
7 h	**76**	**0.37**	**0.25, 0.50**	**3.97**	**6.81**	15.32	**76 (68)**	**59 (53)**
6 h	77	0.36	0.24, 0.49	4.15	7.11	16.00	77 (69)	6 (5)
5 h	80	0.37	0.25, 0.49	3.97	6.81	15.32	79 (70)	2 (2)
4 h	80	0.36	0.25, 0.49	4.15	7.11	16.00	80 (71)	17 (15)
3 h	86	0.39	0.28, 0.51	3.65	6.26	14.08	85 (76)	23 (20)

[a]*None of the children met the criteria for reliabilities of 0.9*
The suggested 'optimum' ICC values for this sample are highlighted. Using the single day ICC value of 0.37, the results of the S-B prophecy formula suggests that for 7 hours of data collection, for 4 days would achieve the desired ICC value of 0.7. This would include 68% of the sample.

The single day ICC values are similar to those reported in earlier studies (Hinkley et al. 2012b). The results of the S-B prophecy formula suggest that when including a weekend day, at least 7 hours of data collection for 4 days is necessary to achieve an ICC value of 0.7. While the ICC values for the inclusion of a weekend day are slightly lower than when 'any' 4 days are used it is not clear of the impact that weekend days have on consistency across days.

To explore the impact of including a weekend day in analysis, reliability coefficients were calculated for data with four weekdays only and for data with three weekdays plus one weekend day with subjects who had at least 7 hours of data (n = 74). Table 8.7 presents the results. Similar to the findings of Penpraze et al. (2006) the reliability coefficient did not change when a weekend day was included. The reliability coefficients in this sample are slightly lower than reported in Penpraze et al. (2006), however, with 4 days of data collection the reliability coefficients are > 70%. The results of the analysis also show that reliability of physical activity measurement depended on the number hours and the number of days of measurement included in the analysis.

Table 8.7: Reliability coefficient (%) of days of data and 95% confidence intervals comparing 4 days of data which includes 1 weekend day and 4 weekdays.

	2 days	3 days	4 days
4 weekdays only	58	68	74
	(45 - 69)	(55 - 77)	(62 - 82)
4 days including 1 weekend day	59	68	74
	(46 - 71)	(56 - 79)	(63 - 83)

8.3.4 Influence of applying different criteria for non-wear time

The influence that the exclusion of different periods of non-wear time has on whether participant's data is included in analysis was explored. Comparison was made between the sample size when the different non-wear time criteria of 10, 20 and 60 minutes of non-wear time were excluded when data were collected over 6, 8 and 10 hours. The results are presented in Table 8.8.

Table 8.8: Influence of non-wear time on number (%) of participants included in analysis.

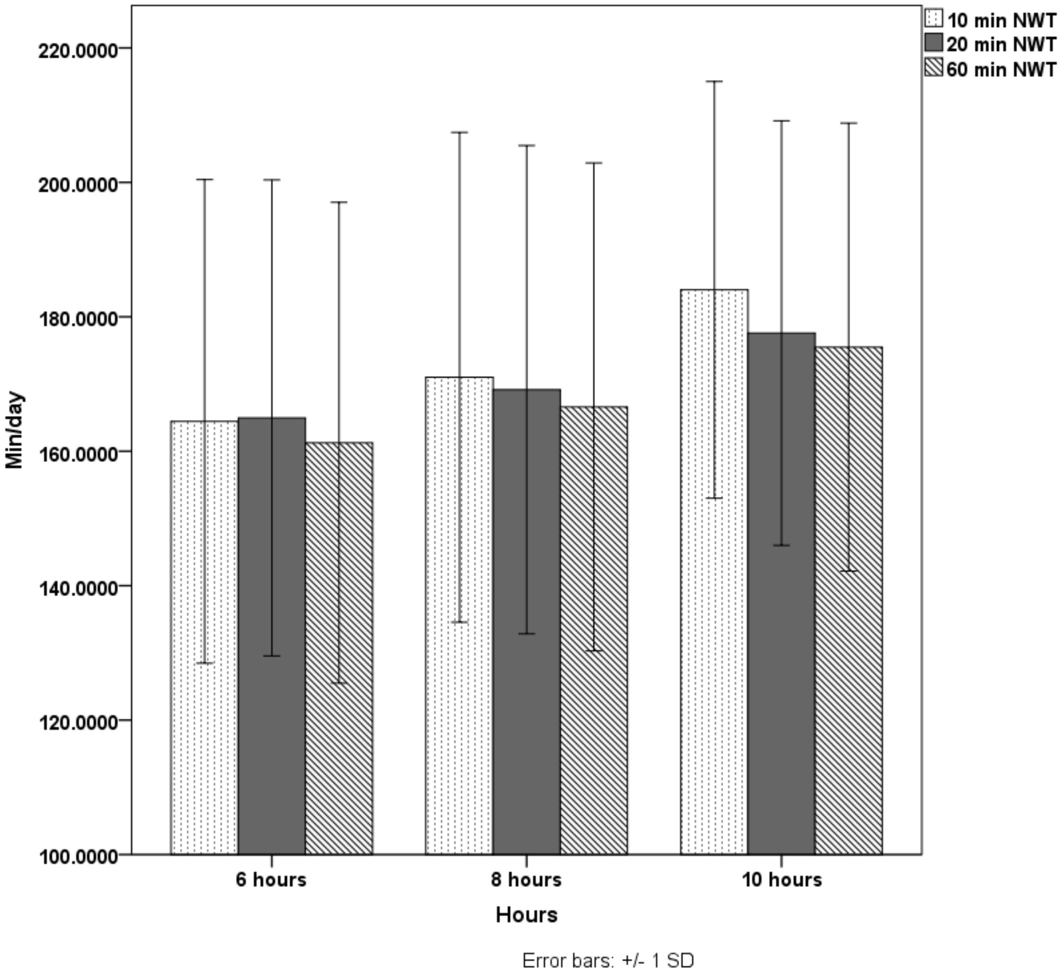
	6 hours			8 hours			10 hours		
Days	**> 10**	**> 20**	**> 60**	**> 10**	**> 20**	**> 60**	**> 10**	**> 20**	**> 60**
1	110 (98%)	110 (98%)	110 (98%)	109 (97%)	109 (97%)	110 (98%)	102(91%)	103 (92%)	103 (92%)
2	107 (96%)	107 (96%)	107 (96%)	105 (94%)	105 (94%)	106 (95%)	92 (82%)	97 (87%)	99 (88%)
3	106 (95%)	106 (95%)	106 (95%)	100 (89%)	100 (89%)	101 (90%)	77 (69%)	83 (74%)	94 (84%)
4	98 (88%)	98 (88%)	100 (89%)	95 (85%)	96 (86%)	97 (87%)	62 (55%)	72 (64%)	81 (72%)
5	97 (87%)	97 (87%)	98 (88%)	85 (76%)	89 (79%)	92 (82%)	46 (41%)	54 (48%)	70 (63%)
6	85 (76%)	85 (76%)	86 (77%)	74 (66%)	80 (71%)	83 (74%)	25 (22%)	37 (33%)	56 (50%)
7	69 (62%)	70 (63%)	72 (64%)	48 (43%)	55 (49%)	61 (54%)	9 (8%)	16 (14%)	24 (21%)
8	5 (4%)	7 (6%)	14 (13%)	2 (2%)	2 (2%)	5 (4%)	0	0	2 (2%)
9	2 (2%)	2 (2%)	3 (3%)	0	0	1 (1%)	0	0	0
10	0	0	0	0	0	0	0	0	0

Boxes highlighted are those where at least 80% of the sample are included

It can be seen in Table 8.8 that the number of hours and days of data collection seems to have an influence on participant inclusion when the different criteria for non-wear time are applied. As the number of hours of data collection increases the 'drop out of participants' increases such that with 10 hours of data collection more than half of participants are excluded when 5 days or more days of data are collected, when the non-wear time criteria of excluding > 10 minutes and > 20 minutes of consecutive zeros are applied. The highlighted boxes are those where more than 80% of participants are included in the sample.

The influence of applying different criteria to exclude non-wear time on estimates of TPA was examined. The data on minutes of TPA was found to be normally distributed ($p > 0.05$) and therefore the mean (SD) minutes of TPA was explored when 10, 20 and 60 minutes of non-wear time were excluded. The results are presented in Figure 8.6.

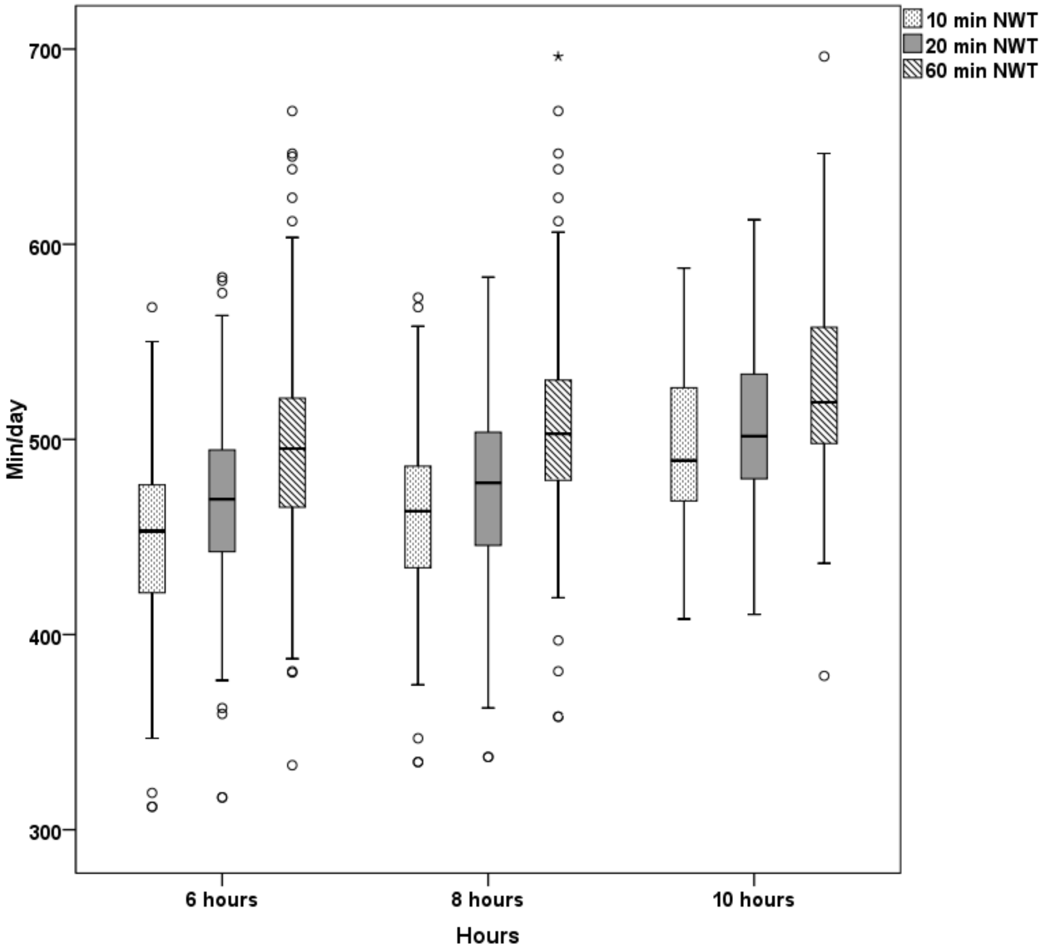
Figure 8.6: Mean (SD) minutes per day spent in total physical activity (TPA) comparing different non-wear time (NWT) criteria.



It can be seen that using shorter periods of non-wear time (10 minutes) result in a greater estimate for the number of minutes of TPA. Differences of 16.1 minutes for the mean time per day spent in TPA was observed when 60 minutes of non-wear time from 6 hours of data (n = 110, mean (SD) min: 161.5 (35.8) min) was compared to 10 minutes of non-wear time with 10 hours of data (n = 102, mean (SD) min: 177.6 (36.9) min).

The influence of non-wear time and hours of data collection on minutes spent in sedentary behaviour was also explored. Analysis showed that this was not normally distributed so the results are presented using box plots with inter-quartile ranges in Figure 8.7.

Figure 8.7: Median (IQR) minutes per day spent in sedentary behaviour comparing non-wear time (NWT) criteria.



It can be seen that for sedentary time, the use of 60 minutes of non-wear time would result in greater estimates of time spent in sedentary behaviour. The median (IQR) minutes spent in sedentary behaviour ranged from 447.1 (61.2) min with the exclusion of 10 min of consecutive zeros from 6 hours of data (n = 110), to 519.3 (53.5) min with the exclusion of 60 min of consecutive zeros from 10 hours of data (n = 103).

8.4 DISCUSSION

8.4.1 Standard measurement day

The results of plotting a 'standardised day' were similar for weekend and week days, with a similar pattern of start and finish times. This is in contrast to the findings of Catellier et al.

(2005) who plotted the start and finish times of adolescent girls, and found a marked difference in the start times and finish times for weekdays in comparison to weekend days. Using a 'standardised' day would mean that the minimum number of hours required for a week day would be 7 hours 14 minutes and for a weekend day 7 hours 48 minutes. In contrast, Catellier et al. (2005) had the minimum number of hours as 11.2 hours for weekdays and 7.2 hours for weekend days. However, the wear times are not necessarily an accurate reflection of times of day when a child or adolescent is awake but only indicate when the accelerometers were worn.

8.4.2 Comparison of weekend and weekdays

In this sample, there was no significant difference between mean cpm for weekdays and weekend days. This is similar to the findings of earlier studies of pre-school children (Oliver et al. 2011). While there was no significant difference in cpm for males between weekend and weekdays, females had significantly higher cpm at weekends compared to weekdays. For the females this was a median difference of 71.3 cpm between weekdays and weekend days. This finding is comparable to the findings of earlier studies where it was argued that this magnitude of difference may not be biologically meaningful (Penpraze et al. 2006). In particular given the cut-points of $\geq$ 3200 cpm (Puyau et al. 2002) used for total physical activity, the small difference in cpm may result in minimal differences in estimates of TPA.

Comparison between the different genders revealed that boys had significantly higher cpm during the week in comparison to girls. The differences in cpm between males and females are reported in earlier studies (Eijkemans et al. 2008; Jackson et al. 2003; Montgomery et al. 2004). There were no differences in cpm for weekend days between males and females.

There were no significant difference in the percentage of time spent in TPA between weekend and weekdays. Children spent a similar percentage of time engaged in TPA during weekend days and weekdays. There were slight differences in TPA between weekend and weekdays whereby children spent up to 1.2 % more time in TPA during weekend days in comparison to weekdays. The pattern of younger children being more active at weekends in comparison to weekdays has been reported in earlier studies and is thought to reverse as a child matures (Treuth et al. 2004b; Trost et al. 2000).

8.4.3 Reliability analysis

The results of the reliability analysis revealed that the single day ICC values were fairly low but despite this they were similar to the findings of earlier studies (Hinkley et al. 2012b; Mattocks et al. 2008). As greater inter-individual variation would result in higher ICC values (Ridley et al. 2009) it is possible that in this sample the inter-individual variation was low which could suggest that in this sample the levels of TPA were similar between participants. While there was some increase in the single day ICC values with increasing number of hours of data collection there was no clear pattern to this. However, similar to the findings by Penpraze et al. (2006) the ICC values in this study were seen to decrease when 10 hours of data collection were used (0.27 for the data including a weekend day and 0.35 for the data from any 4 days). As ICC values are influenced by sample size (Baranowski and de Moor 2000) the differences in the ICC values may in part be attributed to the different sample sizes in each of the hourly groups. It is noteworthy that when data including a weekend day were analysed, the single day ICC values were similar for 3- to 9-hours of data collected (ranging from 0.35 to 0.39) and the differences in the ICC values were small (≤ 0.04). As a result, a similar estimate in the number of days required to achieve a pre-specified level of reliability was seen.

To estimate the number of days of data collection using the S-B prophecy formula it was first necessary to decide on the desired level of reliability. While setting an ICC value of 0.8 may be desirable at the outset, this has an impact on the number of days of data collection required. An ICC can be increased either by decreasing the intra-individual variation of activity between days, by increasing the number of days of measurement or by increasing the sample size (Baranowski and de Moor 2000). In the present study for example, setting a reliability coefficient of 0.8 for data including 1 weekend day meant that in many cases 8 days of data collection would be necessary. One of the consequences of increasing the number of days of data collection is the risk of decreasing the sample size, due to exclusion of participant data which does not reach the minimum number of days criteria. This is particularly concerning when studies involve young children where compliance may be more difficult. It can also impact on the 'completeness' of the data being analysed. Hinkley et al. (2012b) and Mattocks et al. (2008) both argue that an ICC target of 0.7 is an acceptable level of reliability as this level still manages to maximise the power of the study by reducing the number of participants excluded from analysis (Mattocks et al. 2008).

Taking the single day ICC value for 7 hours of data collection, an estimated 3 days of data collection are necessary for ‘any’ 4 days of data and 4 days are necessary when a weekend day is included to achieve an ICC value of 0.7. Selecting 7 hours of data collection as a minimum criteria may help to minimise the possible effects that varying length of day could have on physical activity outcomes (Mattocks et al. 2008). It will also allow for comparison of findings with earlier studies which have typically used 6 to 8 hours of data collection for this age group (Fisher et al. 2005b; Kelly et al. 2006; Kelly et al. 2007).

The single day ICC values for ‘any’ 4 days of data collection were slightly higher (ranging from 0.35 to 0.44) than when a weekend day was included. It is possible that the inclusion of a weekend day increases variability between days and this may be due to differences in the types of activity undertaken during weekend days in comparison to weekdays. To explore the influence of a weekend day, data collected over 4 weekdays were compared with data from 4 days which included a weekend day. It is notable that the inclusion of a weekend day had little effect on the resulting reliability coefficients. This is a similar finding to the study by Penpraze et al. (2006). While the reliability coefficient values in the current study are lower than the Penpraze et al. study (2006) the values were over 70% when 4 days of data was collected for at least 7 hours/day.

An earlier study by Raudsepp and Päll (1998) of older children (8 to 9 y) reported that reliability between weekend and weekdays was lower than when two weekend or two weekdays were compared. The authors argued that the lower reliability may be explained by differences between weekend days where children partake in more unstructured free-time activities and weekdays or school days where activity is more stable. The results of the current study suggest that the reliability between weekend and weekdays is similar. Since younger children in this study were not yet in formal school education it is possible that their activities during the week are less structured and therefore their habitual physical activity behaviour is more similar to activity undertaken during a weekend day. The differences in weekend and weekdays may become more apparent when children start attending school during weekdays.

There are, however, limitations in the present study with the use of ICC values and the S-B prophecy formula. ICC values are constrained to the sample from which they are calculated (Baranowski and de Moor 2000), and the magnitude of the intra- and inter-individual variances in physical activity are sample-specific (Ridley et al. 2009). In addition, the S-B prophecy formula assumes that the ICC remains the same when additional monitoring days are added (Ridley et al. 2009).

8.4.4 Influence of applying different criteria for non-wear time

The analysis of the influence of non-wear time on participant involvement suggested that both the hours of data collection and the non-wear time periods influenced the sample size. With 10 hours, the sample size fell below 80% with more than 2 days of data collected when the non-wear time criteria to exclude 10 and 20 minutes of consecutive zeros were applied. With 6 and 8 hours, the periods of non-wear time did not seem to greatly influence the exclusion of participants and a greater influence came from the criteria set for the hours of data collection.

The effect of applying different non-wear time criteria and hours of data collection had an influence on estimates of TPA. It should be noted that the sample sizes of these groups were not the same, and therefore this analysis includes data from different participants which may explain some of the differences in outcome. To explore the influence of non-wear time on the same group of participants, this analysis was repeated for participants who had complete data for 10 hours collected over 4 days (n = 62). The results are presented in Table 8.9. The differences between time spent in TPA were less apparent when the criteria for 10 minutes and 60 minutes of non-wear time were applied to the data (approximately 3 minutes). There was however, a more apparent difference for time spent in sedentary time (approximately 37 minutes).

Table 8.9: Median (IQR) minutes of total physical activity and sedentary behaviour per day comparing non-wear time criteria applied to 'complete' participant data.

	Exclusion of periods of non-wear time (n = 62)		
	10 min	**20 min**	**60 min**
TPA	181.87 (32.57)	179.18 (33.42)	176.51 (32.97)
Sed	491.48 (40.55)	503.90 (42.96)	528.80 (53.57)

Complete data: participants with 4 days of 10 hours of data; Sed: sedentary behaviour; TPA: total physical activity.

The longer the non-wear time excluded (e.g. 60 minutes of non-wear time), the greater the time consequently classified as sedentary. Similar findings have been reported in a study by Evenson et al. (2009) of post partum women (n = 182) whereby excluding longer periods of non-wear time (60 versus 20 minutes of consecutive zeros) led to an increase in estimated sedentary time and a decrease in percentage time classified as MVPA. In a study involving adolescent girls (n = 1,348, 11 - 13 years), Toftager et al. (2012) reported that the longer the period of non-wear time excluded, the lower apparent physical activity levels (cpm). It is possible that as the number of consecutive zeros used to define non-wear time increases, there will be more zeros left within the data to be included in the analysis and therefore a greater amount of time is classified as sedentary. For example, as there will likely be fewer sustained episodes of 60 minutes of consecutive zeros, in comparison to 10 minutes of consecutive zeros in the data, using the 60 minute of non-wear time criteria would result in less of the zero count data being excluded.

While the analysis of this data does not give an indication of which criteria are the optimal to use for non-wear time, the findings do suggest the criteria applied will have an influence on the overall outcome and this is complicated by the minimal number of hours selected. Researchers have argued that it is likely that adults will remain still for longer periods than children and it may be appropriate to select thresholds with longer durations of consecutive zeros with adults (Mâsse et al. 2005). In reviewing the evidence from the literature on bouts of non-wear time, the only studies to have investigated this in children are by Esliger et al. (2005) and Alhassan et al. (2008) and both these studies suggest that excluding data with greater than 20 minutes of consecutive zeros may be a reasonable approach to adopt. It remains to be seen whether age-specific criteria for non-wear time are necessary and further investigation is warranted.

8.4.5 Conclusion

The results of the present study suggest that there were no differences in accelerometer cpm between weekend and weekdays. However, some gender differences were apparent. There were also no significant differences in percentage time spent in TPA and the differences in TPA between weekend and weekdays was no more than 1.2%. The pattern of wear time for the accelerometers was similar between weekend and weekdays which may support the hypothesis that younger children are active for a similar amount of time at the weekend compared to weekdays and that this differs from what has been reported in studies of adolescents (Catellier et al. 2005). The inclusion of a weekend day had minimal influence on reliability coefficient values and hence the inclusion of a weekend day is not necessary for this sample. While the ICC values were similar for all hours of data collection, 7 hours of data collection had the highest ICC value (0.44) and it is estimated that 3 days of data collection would achieve an ICC value of 0.7. The application of different non-wear time criteria, to exclude periods of consecutive zeros, did impact on estimates of time spent in TPA and sedentary behaviour. Excluding 60 minutes of non-wear time resulted in more time classified as sedentary and less time classified as TPA in comparison to 10 minutes of non-wear time. This has to be considered when comparing results of different studies.

CHAPTER 9 : GENERAL DISCUSSION

9.1 INTRODUCTION

This thesis set out to make an original contribution to knowledge in relation to how accelerometers are used to accurately measure free-living physical activity and sedentary behaviour of pre-school children. As outlined in the introduction (Chapter 1), accurate measurement is vital to gain a deeper understanding of the relationship between physical activity, sedentary behaviour and health. It is particularly important for determining the success of interventions aimed at the promotion of physical activity or reduction of time spent in sedentary behaviour. It also supports the surveillance of physical activity and sedentary behaviour in populations including tracking of behaviours over time. Physical activity behaviour in young children is believed to be distinct to physical activity behaviour seen in older children, whereby younger children tend to engage in more sporadic bursts of high intensity physical activity (Bailey et al. 1995). However, the true characteristics of physical activity and sedentary behaviour in the pre-school population are not well understood. Moreover, varying degrees of methodological uncertainty exist with regard to accurate accelerometer-based quantification of physical activity and sedentary behaviour of young children (Cliff et al. 2009b). This thesis is based on six empirical studies, each one addressing a specific and discrete methodological question. The questions upon which this thesis are based and which relate to the accuracy of measurement of physical activity in pre-school children are as follows:

- What are the implications of shorter epochs on estimates of physical activity and sedentary behaviour in pre-school children (i.e. is there an epoch effect)? (Chapter 3).
- Which epoch is most accurate for measurement of physical activity in pre-school children? (Chapter 4).
- Are there advantages of using triaxial over uniaxial accelerometry to measure physical activity in pre-school children? (Chapter 5).
- Which Actigraph accelerometry cut-points are most accurate for pre-school children? (Chapter 6).
- Are different generations of Actigraph accelerometers comparable when used with pre-school children? (Chapter 7).
- What is the recommended wear time to provide a reliable estimate of habitual physical activity and sedentary behaviour in pre-school children? (Chapter 8).

Table 9.1 summarises the research questions posed in this thesis and the key findings.

The purpose of this final chapter is to consider and synthesise the findings from the six empirical study chapters in greater depth and to evaluate these finding in the context of the existing literature. This chapter will also consider the limitations of the studies and discuss the implications of these empirical observations for future research. Finally, this chapter will make recommendations for investigators and policy makers regarding measurement of physical activity and sedentary behaviour in pre-school children.

Table 9.1: Summary of key findings from thesis.

Research question/Study	Key findings
Is there an epoch effect? (Chapter 3)	• There was an epoch effect with 15-s epochs resulting in significantly more minutes classified as MVPA in comparison to 60-s epochs. • There was a moderate effect size (r = -0.32) for this average difference of 11 min/day.
Which epoch is most accurate? (Chapter 4)	• Good absolute agreement between accelerometry estimates of MVPA at 15-s epochs and the CARS criterion measure, suggesting that 15-s epochs may be more accurate than 60-s epochs for detecting time spent in MVPA.
Are there advantages to using triaxial accelerometry? (Chapter 5)	• No advantage of using the RT3 triaxial accelerometer over a uniaxial accelerometer for studies of pre-school children for assessment of either relative or absolute amounts of physical activity. • The Sun et al. (2008) cut-point for the RT3 for light jog (780 counts/15 s) provides a reasonable estimate of MVPA against the CARS criterion measure.
Which Actigraph accelerometry cut-points are most accurate? (Chapter 6)	• The Puyau et al. (2002) cut-points provided estimates of time spent in light intensity physical activity, MVPA and sedentary behaviour that had good absolute agreement with the CARS criterion measure. • The Sirard et al. (2005) cut-points had good agreement with the CARS for MVPA, but overestimated time spent in sedentary behaviour and underestimated time spent in light intensity activity.
Are different generations of Actigraph accelerometer comparable? (Chapter 7)	• Good agreement between 7164 output and GT1M output for total physical activity when a correction factor is applied to the GT1M data (7164= GT1M/0.91). • Applying a correction factor may partially resolve the issue of the output from the GT1M having lower cpm than the 7164 model.
What is the recommended wear time to provide a reliable estimate of habitual physical activity and sedentary behaviour in pre-school children? (Chapter 8)	• Small differences (< 1.2% difference over a day) between estimates of percentage time spent in total physical activity between weekend and weekdays. • Three days of data collection for a minimum of 7 hours provided adequate reliability (ICC = 0.7). Inclusion of a weekend day did not influence reliability. • Different criteria for 'non-wear time' (e.g. exclusion of 10, 20 and 60 minutes of consecutive zeros) had an impact on estimates of time spent in sedentary behaviour and total physical activity.

CARS: Children's Activity Rating Scale; MVPA: moderate-to-vigorous physical activity

9.2 EPOCH EFFECT AND WHICH EPOCH IS MOST ACCURATE?

It has been argued that shorter epochs (< 60-s epochs) may be more accurate than longer epochs (e.g. 60-s epochs) in capturing high intensity physical activity, particularly in young children (Trost et al. 2005). The first epoch study (Chapter 3) which was published in 2008, set out to investigate the implications of using different accelerometry epoch lengths to measure time spent in MVPA from 32 free-living pre-school children (5 to 6 y, mean (SD) = 5.9 (0.7) y) for data collected over a 7 - 10 day period. At the time of publication, the recommendations for health were for pre-school children to engage in 60 minutes of MVPA per day (Scottish Executive 2003). The results suggested that there was a significant difference in estimated daily time spent in MVPA with 15-s epochs, resulting in approximately 11 minutes more time in MVPA than the 60-s epoch and that there was a moderate effect size for this difference ((r = -0.32) see Chapter 3, Table 3.3, p. 96).

While there were differences in the estimates of the median number of minutes of time per day spent in moderate intensity physical activity, for which 15-s epochs resulted in just under 10 minutes more time in comparison to 60-s epochs, there was only a small difference in the estimates of median daily minutes of vigorous intensity activity between the epochs (1.3 minutes, between 15- and 60-s epochs). There was a more apparent difference for LPA activity, where 60-s epochs resulted in 28 minutes more time being classified as LPA in contrast to the shorter epoch of 15 s. The results of the current study may partially support the hypothesis that longer epochs can lead to a 'smoothing' effect, whereby the averaging of episodes of higher intensity activity with low intensity activity in the same epoch could result in an underestimation of true levels of higher intensity activity (Trost et al. 2005). However, while an earlier study by Nilsson et al. (2002) reported that longer epochs resulted in 'hard' and 'very hard' physical activity being re-classified as 'moderate' intensity physical activity, in the current study longer epochs resulted in more minutes classified as LPA.

What does not seem to concur with the pattern for less time classified at higher intensity, is that the longer 60-s epoch resulted in an estimated 3 minutes less time spent in sedentary behaviour compared to the 15-s epoch. The only other study to date which has reported on the influence of epoch length on sedentary behaviour is by Ojiambo et al. (2011) who reported a similar pattern with 86 school aged children (mean (SD) age: 7 (2) y) measured over 6 days, when the Puyau et al. (2002) cut-points were applied. However, this result may be reflective of the cut-points applied to differentiate between sedentary behaviour and LPA. In the study by Ojiambo et al. (2011) the interaction between epochs and cut-points was illustrated. The authors reported that the application of the Reilly et al. (2003) and the Sirard et al. (2005) cut-points to the accelerometry data resulted in more time classified as sedentary behaviour with longer epochs. This is in direct contrast to what was observed when the Puyau et al. (2002) cut-points were applied. This might be due to the lower cut-point threshold for the Puyau et al. (2002) (< 800 cpm) for classifying sedentary behaviour, which could result in misclassification of some of the sedentary behaviour time as time spent in LPA. In contrast, the higher thresholds for sedentary behaviour for the Reilly et al. (2003) cut-points (< 1100 cpm) and the Sirard et al. (2005) cut-points (< 1207,< 1452 and < 1592 cpm for 3-, 4- and 5-year-old children respectively) may result in more time classified as sedentary behaviour instead of LPA. However, it should be noted that in the first epoch study of this thesis, the 7164 Actigraph accelerometer was used to collect data and, as was revealed in Chapter 7, the output from the GT1M is not necessarily directly comparable with the output of the 7164 accelerometer. So, while the higher cut-points for sedentary behaviour may be suitable for the 7164 accelerometer, they are not necessarily accurate for the GT1M accelerometer. Moreover, in the Ojiambo et al. (2011) study the ActiTrainer (Actigraph, Fort Walton Beach, FL, USA) accelerometer was used and it is not clear if this accelerometer model is comparable with other Actigraph models.

The ability to make an accurate distinction between sedentary behaviour and LPA has become particularly important in recent years given the change to the recommendations for health, in that pre-school children should engage in 180 minutes of 'total' physical activity (combining LPA and MVPA activity) per day (Australian Government, Department of Health and Ageing 2009; Canadian Society of Exercise Physiology 2012; Department of Health, Physical Activity, Health Improvement and Protection 2011). Given the updated recommendations for health, the data were revisited to investigate the influence that epochs had on time spent in TPA when the Puyau et al. (2002) cut-points were applied (Table 9.2).

Table 9.2: Median (IQR) minutes of total physical activity (TPA) with Puyau et al. (2002) cut-points.

	Epoch (seconds)		
	15 s	**30 s**	**60 s**
TPA	174.3 (68)	185.8 (83)	192.0 (95)

TPA: total physical activity.

The results in Table 9.2 suggest that the longer epoch resulted in a significantly greater amount of time being classified as TPA (z = -14.1, p = 0.00, r = -0.6). This is important, as researchers need to be aware of the influence that different epochs have on outcomes. It is not necessarily as straightforward as shorter epochs resulting in more time spent in higher intensity activity, because shorter epochs may also influence the classification of time spent in low intensity physical activity and time spent in sedentary behaviour.

Since publication of the findings from Chapter 3 in 2008 (Reilly et al. 2008), the epoch effect, which can impact on estimates of time spent in MVPA, has been reported in other published studies with school-aged and pre-school children (Edwardson and Gorely 2010; Mahar et al. 2008; Ojiambo et al. 2011; Vale et al. 2009). However, these earlier studies with pre-school children have been limited to data collected from school days (Vale et al. 2009), or have been based on data collected over one day (Mahar et al. 2008). The novel contribution of the current study is that it is the only study to date which has investigated epoch effect in pre-school children over an extended period of time (7 - 10 days).

In the second epoch study of this thesis (Chapter 4), the accuracies of the different epochs were compared against a criterion measure using the CARS (1990) direct observation tool with 31 pre-school children (mean (SD) age: 4.4 (0.8) y) during their 1 hour of time-tabled outdoor play time. This study provided evidence to support the widespread perception that shorter epochs are essential to accurately measure MVPA in young children. The results revealed that estimates from 15-s epochs had good absolute agreement with the CARS criterion measure (mean difference (LOA): 0.8 (-6.22 to 7.8) min), in comparison to estimates made with 60-s epoch, which had a greater mean difference from the criterion measure (mean difference (LOA): 3.2 (-6.7 to 13.2) min).

The study also revealed that shorter epochs of 1 and 5 s resulted in higher levels of MVPA than the 15-s epochs (Chapter 4, Table 4.2). A limitation of this study was that the CARS criterion measure in this study was undertaken in 15-s observation periods. Consequently it is not clear whether shorter epochs would offer greater accuracy, and it is possible that 15-s epochs may be underestimating time in MVPA. It is plausible to speculate that this may be the case and authors have indeed argued that shorter epochs would offer greater sensitivity to detect changes in physical activity intensity (Ayabe et al. 2013). While this was not possible in the early days of accelerometry, with the advanced memory storage capabilities of currently available accelerometers, the collection of data using shorter epochs over extended periods of time is now possible (Rowlands 2007).

In the Chapter 4 epoch study, it became apparent that use of the Sirard et al. (2005) cut-points led to significant differences between the estimates of sedentary behaviour, LPA and TPA and the CARS for both 15- and 60-s epochs (p < 0.001). Interestingly, in this study both epochs produced an 'underestimation' of the TPA and the accelerometry estimate of sedentary behaviour was much higher than the CARS estimate. This could suggest that the higher cut-points by Sirard et al. (2005) may be misclassifying time spent in LPA as time spent in sedentary behaviour.

In conclusion, the use of longer epochs resulted in a greater amount of time being classified as LPA, with shorter epochs resulting in more time classified as moderate intensity physical activity. Very little time was seen to be spent in vigorous activity and, therefore, any differences between different epochs for vigorous activity were less apparent. There was good absolute agreement between accelerometry estimates of MVPA at 15-s epochs and the criterion measure of direct observation, suggesting that the 15-s epochs may be more accurate than 60-s epochs for detecting time spent in MVPA. There were, however, large limits of agreement, so while the accuracy may be adequate at a group level, this may not be the case for individual classification of physical activity intensity.

Table 9.3 presents the summary of the key finding for Chapter 3 and Chapter 4.

Table 9.3: Summary of key findings from Chapter 3 and 4.

Is there an epoch effect? (Chapter 3)	• Shorter epochs (15 s) resulted in significantly more time classified as MVPA than 60-s epochs.
Which epochs are most accurate? **(**Chapter 4)	• Good absolute agreement between accelerometry estimates of MVPA at 15-s epochs and the CARS criterion measure, suggesting that 15-s epochs may offer greater accuracy than 60-s epochs for detecting time spent in MVPA.

9.3 ARE THERE ADVANTAGES TO USING TRIAXIAL ACCELEROMETRY?

In Chapter 5, the aim was to investigate whether there were advantages to using triaxial over uniaxial accelerometers to measure physical activity in pre-school children. To allow comparison between the models, the GT1M uniaxial accelerometer and the RT3 triaxial accelerometers were worn simultaneously by 31 pre-school children (mean (SD) age: 4.4 (0.8) y) while they engaged in 1 hour of free-play. The accelerometry count output from both monitors was compared against direct observation using the CARS (Puhl et al. 1990) scale as the criterion method for time spent at different intensities of physical activity and sedentary behaviour. The results of the study suggest that there was no evidence that the triaxial RT3 accelerometer was more accurate than the GT1M accelerometer for estimating absolute amount of time spent in MVPA. While there was a positive correlation between the output from the RT3 and GT1M models, accounting for almost 50% of the shared variance ($r = 0.72$, $p < 0.001$), there were marked differences between the accelerometer models in estimates of time spent in MVPA, which varied according to the cut-point threshold applied to the data (Chapter 5, Table 5.2). In addition, there was an overestimation of time spent in MVPA for all cut-points for the RT3 when compared against the CARS (Puhl et al. 1990) criterion method (Chapter 5, Table 5.2).

At the time of publication of Chapter 5 in 2012 (Hislop et al. 2012b), there were no published cut-point thresholds for the RT3 for pre-school population. However, later in 2012, Adolph et al. reported on their calibration study in which they undertook whole room direct calorimetry with 64 pre-school children (mean (SD) age: 4.5 (0.8) y) over a 3-hour period. They reported that the accelerometer counts between the RT3 and the uniaxial models (Actical and Actiheart) were positively correlated with each other ($r = 0.8 - 0.95$, $p = 0.001$) and that the overall classification accuracy (71% and 72%) was similar for the RT3 (73%) and the uniaxial accelerometers respectively (Adolph et al. 2012). The authors identified a cut-point for MVPA of 1400 cpm, or 350 counts/15 s, which is lower than the 413 counts/15 s identified by Sun et al. (2008) as 'walking relaxed'. Given that in the current study, the Sun et al. (2008) cut-point for walking relaxed resulted in an overestimation of MVPA (19.3 min for the Sun et al. (2008) cut-points compared to 4.8 min for the CARS), lowering the threshold for MVPA further would result in a greater difference between the accelerometer estimated time in MVPA and the CARS (Puhl et al. 1990). However, Adolph et al. (2012) also identified cut-points for specific 'moderate' intensity activities such as ball play and dancing/aerobics (468 and 590 counts/15 s) and a fast translocation defined as 'jog' (691.5 counts/15 s). The cut-points for these 'moderate' intensity activities are closer to the Sun et al. (2008) cut-point of 780 counts/15 s for a light jog, which in the current study (Chapter 5) resulted in an estimate for MVPA which was most similar to the CARS criterion measure of MVPA. Despite this, Adolph et al. (2012) reported a good correlation between RT3 accelerometer counts from the vector magnitude measure and the CARS (r = 0.74).

One explanation for the differences in the cut-point thresholds between the current study and the study by Adolph et al. (2012) relates to the differences in the calibration processes used in the studies. In the Adolph et al. (2012) study, the thresholds were based on established heart rate (HR) cut-offs to predict Activity Energy Expenditure (AEE), and in the current study a behavioural approach to quantify physical activity behaviour was used. While the current study is limited in so far as it is not possible to predict AEE from the CARS approach, there are also concerns with using HR as a proxy for physical activity, as this can be affected by factors other than physical activity, such as emotions, surroundings, dehydration and temperature (Muller and Bosy-Westphal 2003). It has also been argued that, rather than being a measure of physical activity, HR is more appropriately used to provide an indirect approximation of energy expenditure (Oliver et al. 2007b).

In conclusion, the results of the current study suggest that there is no advantage to using the RT3 triaxial accelerometer over a uniaxial accelerometer in studies of pre-school children for assessment of either relative or absolute amounts of physical activity. Results from this study suggest that, if behavioural measures of physical activity are of interest, then a higher cut-point threshold for the RT3, similar to the Sun et al. (2008) cut-point for light jog (780 counts/15 s), would provide a more reasonable estimate of MVPA. Table 9.4 provides a summary of the key findings from Chapter 5.

Table 9.4: Summary of key findings from Chapter 5.

Are there advantages to using triaxial accelerometry? (Chapter 5)	• No advantage to using the RT3 triaxial accelerometer over a uniaxial accelerometer for studies of pre-school children, for assessment of either relative or absolute amounts of physical activity. • The Sun et al. (2008) cut-point for the RT3 for light jog (780 counts/15 s), provides a reasonable estimate of MVPA against the CARS criterion measure.

9.4 WHICH CUT-POINTS ARE MOST ACCURATE?

In the study presented in Chapter 6, the principle aim was to validate Actigraph accelerometry cut-points for estimating physical activity and sedentary behaviour in pre-school children during free-play. In this study, estimates of time spent at different intensities (sedentary behaviour, LPA and MVPA) were compared when different accelerometry cut-points were applied to data collected from 31 pre-school children (mean (SD) age: 4.4 (0.8) y) during 1 hour of free-play. The accuracy of the cut-points was determined using the CARS (Puhl et al. 1990) criterion measure of physical activity in the same population.

The result of this study highlighted the marked differences in outcome when different cut-points were applied (Chapter 6, Table 6.2). The impact that these differences have on apparent levels of MVPA and time spent in sedentary behaviour have been highlighted in a number of earlier studies (Guinhouya et al. 2006; Mota et al. 2007; Reilly et al. 2008). In the Reilly et al. (2008) study, for example, estimates of MVPA ranged from 28 min MVPA per day with the Puyau et al. (2002) cut-points to 226 min (nearly 4 hours) of MVPA with the Freedson et al. (1997) cut-points. However, while these studies have described the differences in outcome with different cut-points, in contrast to the current study, they have not addressed the issue of the accuracy of various cut-points.

The study presented in Chapter 6 indicated that the Pate et al. (2006) and the Van Cauwenberghe et al. (2011) cut-points significantly overestimated minutes of MVPA in pre-school children compared against a direct observation method. Use of the Puyau et al. (2002) and Sirard et al. (2005) cut-points produced estimates of MVPA (median minutes of: 4.3 min and 4.8 min respectively), which were not significantly different from the criterion measure (CARS: 4.8 min). The Puyau et al. (2002) thresholds for sedentary behaviour, LPA and for TPA were the only other cut-points which provided estimates that were not significantly different from the criterion measure and provided agreement at a group level with the criterion measure.

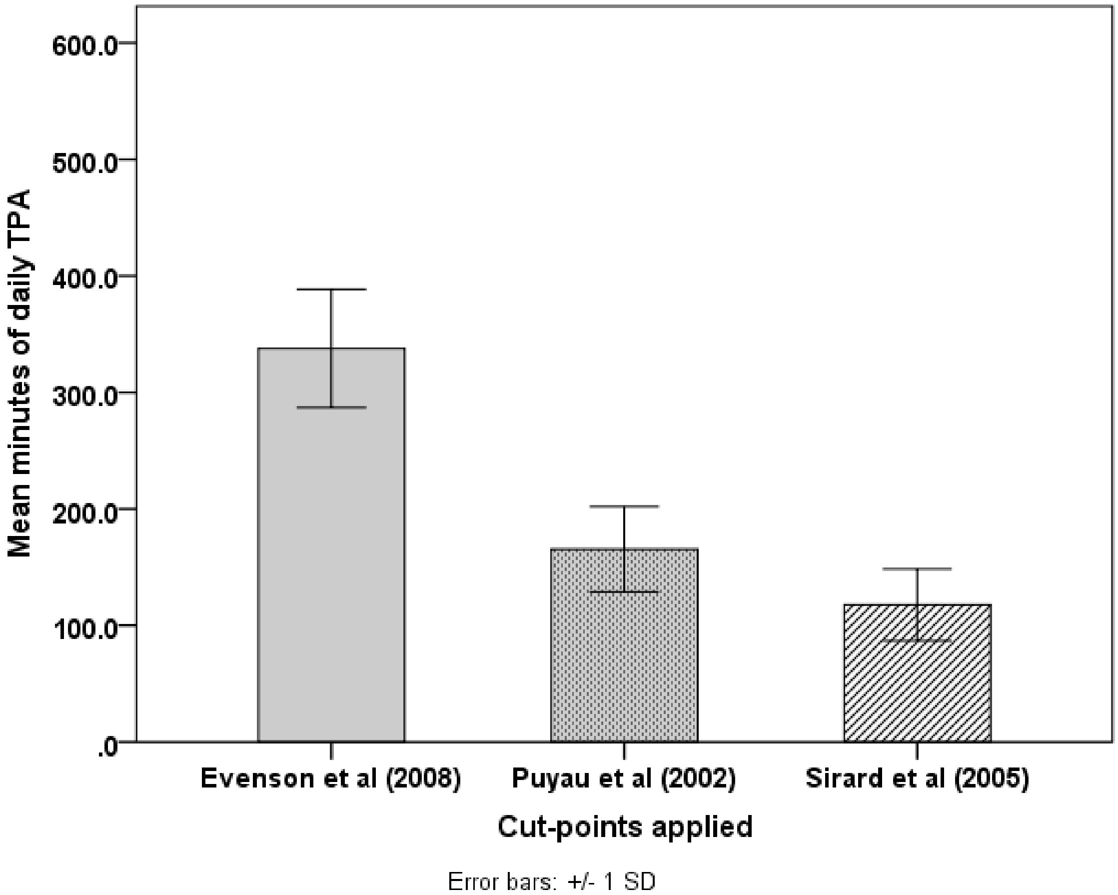
The reason that different cut-points have been developed might again be explained by the differences in calibration methods adopted to validate the thresholds. Pate et al. (2006) and Evenson et al. (2008) used indirect calorimetry as a measure of energy expenditure, however, in the current study and the study by Van Cauwenberghe et al. (2011) the CARS direct observation tool was used as the criterion method. The differences between the current study and the Van Cawenberghe et al. (2011) study may in part be due to differences in interpretation of the scoring system for the CARS tool. In the current study MVPA was classified as activity at CARS Levels 4 and 5 which were grouped together. In contrast, Van Cauwenberghe et al. (2011) classified moderate activity as averaged CARS scores of between 3.1 and 4 and vigorous as being averaged scores of between 4.1 and 5. The lower CARS scores used by Van Cauwenberghe et al. (2011) for moderate intensity in their study meant that they included activities with an average energy expenditure of less than three times an individual's resting energy expenditure, which in the original CARS study was considered to be LPA (Puhl et al. 1990). This included activities described as being 'translocation' (slow/easy), such as slow walking (Puhl et al. 1990). The current study used the higher average CARS score of more than or equal to four to define MVPA, since this represented an average energy expenditure of more than three times individual resting energy expenditure according to the original CARS study (Puhl et al. 1990). In addition, the approach outlined in Sirard et al. (2005) involved coding the observations over a 15-s period instead of the second-by-second direct observation coding used by Van Cauwenberghe et al. (2011). Finally, another explanation may be that the Van Cauwenberghe et al. (2011) study made use of indoor play, while the current study involved outdoor play, it is possible that the patterns and intensity of activity undertaken in these environments are different.

It is interesting to note that use of the lower threshold for sedentary behaviour recommended by Puyau et al. (2002) (< 200 counts/15 s) was not significantly different (z = -2.5, p = 0.01, r = -0.3) from the CARS criterion method (Chapter 6, Table 6.2). Cut-points employing a higher threshold for sedentary behaviour characterisation, such as those advocated by Sirard et al. (2005) (< 301 counts/15 s for 3 year olds to < 398 counts/15 s for 5 year olds) and Van Cauwenberghe et al. (2011) (< 373 counts/15 s), which both used the CARS criterion measure, resulted in an overestimation of time spent in sedentary behaviour. While the higher threshold for sedentary behaviour might be due to the Sirard et al. (2005) calibration study using the 7164 Actigraph model, which as observed in Chapter 7 may be recording 9% higher than the GT1M, this does not explain the difference in thresholds with the Van Cauwenberghe et al. (2011) study which used the GT1M accelerometer. The reason for the differences is more likely explained by the cut-points being calibrated against different sedentary activities. In the Puyau et al. (1990) study, the sedentary activities were all undertaken in sitting, while in the Sirard et al. (2005) and the Van Cauwenberghe et al. (2011) study, stationary activities, which included both sitting and standing, were classified as sedentary activities. This will be discussed further in the limitations section of this chapter.

Recent years have seen the emergence of an argument for the use of a lower threshold ($\leq$ 100cpm or $\leq$ 25 counts/15 s) (Evenson et al. 2008) to classify time spent in sedentary behaviour (Trost et al. 2011). However, the application of this threshold has yet to be validated in pre-school children. The threshold of $\leq$ 25 counts/15 s for classifying sedentary behaviour, validated by Evenson et al. (2008), in a study of 33 school-aged children (mean (SD) age: 7.3(1.1) y), is considerably lower than the cut-points proposed by Puyau et al. (2002) and Sirard et al. (2005). Figure 9.1 and Figure 9.2 illustrate the differences in outcome for the average time spent in sedentary behaviour and total physical activity when the Evenson et al. (2008) Puyau et al. (2002) and the Sirard et al. (2005) cut-points are applied to the same set of free-living data collected from 104 children with at least 7 hours of data for 3 days. The data were tested for normality using the Kolmogorov-Smirnov test and data on sedentary behaviour were not normally distributed ($p < 0.05$), while the data on TPA were normally distributed ($p > 0.05$) (Appendix IX, Appendix Table IX.i).

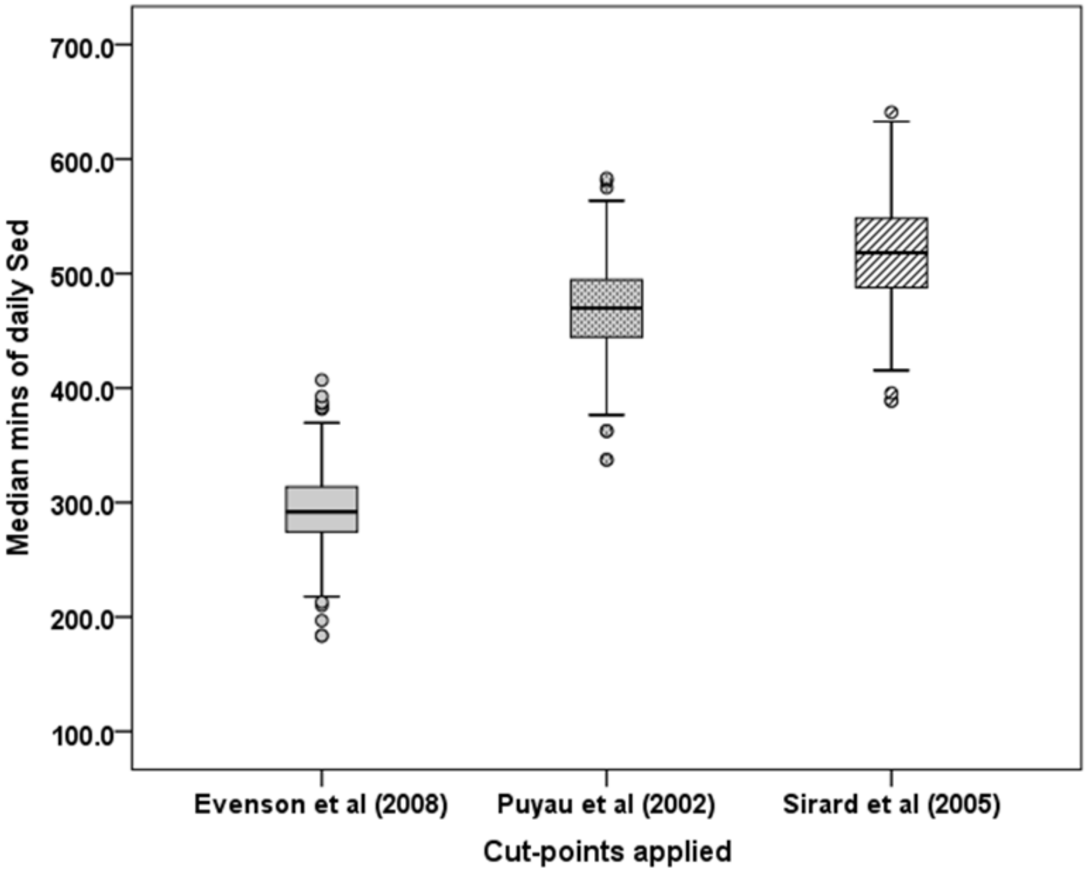
When the Evenson et al. (2008) cut-points are applied 100% of the sample would be seen to be meeting the recommendation of 180 minutes of total physical activity per day (mean (SD) min/day: 337.8 (50.8) min). With the application of the Puyau et al. (2002) cut-points, 31% of the sample met the recommendations (mean (SD) min/day: 165.4 (36.7) min) and with Sirard et al. (2005) cut-points, only 3% of the children in the sample achieved the 180 minutes per day recommendation (mean min/day (SD) = 117.7 (30.8) min). The differences in output that can result from applying different cut-points to process the accelerometry data are of concern and will have major implications for researchers and policy-makers in their attempts to determine whether pre-school children are sufficiently active for health or not.

Figure 9.1: Comparison of mean (SD) minutes per day using Evenson et al (2008), Puyau et al (2002) and Sirard et al (2005) cut-points for total physical activity (n=104).



TPA: total physical activity.

Figure 9.2: Comparison of median (IQR) minutes per day using Evenson et al. (2008), Puyau et al. (2002) and Sirard et al. (2005) cut-points for sedentary behaviour (n=104).



Sed: sedentary behaviour.

As discussed above, differences in the calibration methods in the original studies might be one explanation for the different cut-points proposed. The Evenson et al. (2008) study involved a structured protocol of activities using indirect calorimetry as a criterion measure for the 7164 accelerometer. In contrast, the current study used a behavioural approach as the criterion method while children engaged in unstructured outdoor free-play activities. It is possible that the nature of free-play means that the episodes of physical activity are likely to be more sporadic and intermittent than during a structured protocol. Similar to what was discussed earlier, the Evenson et al. (2008) study only included sitting activities in their protocol for ‘sedentary behaviour’, while the CARS method includes stationary standing as a ‘sedentary behaviour’ (Puhl et al. 1990). Currently there is no clear consensus in the literature relating to the behaviours and, in particular, stationary standing, which might contribute to sedentary behaviour (De Decker et al. 2013). This topic will be discussed more fully in the limitations section of this chapter. Finally, the differences in the output between

GT1M and 7164 accelerometers also raises questions about the validity of cut-points developed for the 7164 accelerometer being applied to output from the GT1M accelerometer.

In conclusion, application of the Puyau et al. (2002) cut-points provided estimates of time in sedentary, LPA and MVPA that had good absolute agreement with the CARS criterion measure. The Sirard et al. (2005) cut-points for MVPA had better agreement with the CARS, but the cut-points for sedentary behaviour and LPA resulted in an overestimation of time spent in sedentary behaviour and an underestimation of time spent in light intensity physical activity. Table 9.5 summarises the key findings from Chapter 6.

Table 9.5: Summary of key findings from Chapter 6.

Which cut-points are most accurate? (Chapter 6)	• The Puyau et al. (2002) cut-points provided estimates of time in sedentary, LPA and MVPA that had good absolute agreement with the CARS criterion measure. • The Sirard et al. (2005) cut-points had good agreement with the CARS for MVPA, but overestimated time spent in sedentary behaviour and underestimated time spent in LPA.

9.5 ARE DIFFERENT GENERATIONS OF ACTIGRAPH ACCELEROMETERS COMPARABLE?

The aim of the study outlined in Chapter 7 was to compare different generations of Actigraph accelerometers. In the first study a mechanical set up was used to simultaneously oscillate the 7164 and the GT1M Actigraph accelerometers through a range of accelerations. During a second study, the 7164 and GT1M Actigraph accelerometers were compared while worn simultaneously by 23 pre-school children (mean (SD) age: 4.3 (0.8) y) during 1 hour of free-play. The estimates for time spent at different intensities (sedentary behaviour, LPA and MVPA) were compared between the different models and against the CARS (Puhl et al. 1990) as the criterion method.

The results suggest that there was a strong correlation in counts/s between models (r = 0.95, p < 0.01), with a mean difference (LOA) of 3.2 (0.16 to 6.22) counts/s. During free-play the 7164 recorded higher cpm than the GT1M which, although the models were positively correlated (r = 0.70, p < 0.01), resulted in a mean difference of 632 cpm with wide limits of agreement (LOA: -1103.7 to 2092.7). When considering the impact that these differences had on time spent at different intensities, the GT1M recorded significantly less time in MVPA and more time in LPA in comparison with the 7164 accelerometer (Chapter 7, Table 7.3 and Table 7.4). Application of a correction factor to the GT1M data (7164 = GT1M/0.91) recommended by Corder et al. (2007) did not resolve the differences for time spent in MVPA. However, applying the correction factor to the GT1M data would allow comparison with the output from the 7164 accelerometer for time spent in TPA. This was appropriate for the Sirard et al. (2005), Puyau et al. (2002) and the Van Cauwenberghe et al. (2011) cut-points. The results therefore suggest that a correction factor may be of use when comparing the GT1M data with the 7164 accelerometer for TPA.

Given the recent support for the Evenson et al. (2008) cut-points for sedentary behaviour, the data were revisited to consider median time (IQR) spent in sedentary and TPA comparing the 7164 and the GT1M Actigraph accelerometers. The results of the comparison between the outcomes are presented in Table 9.6.

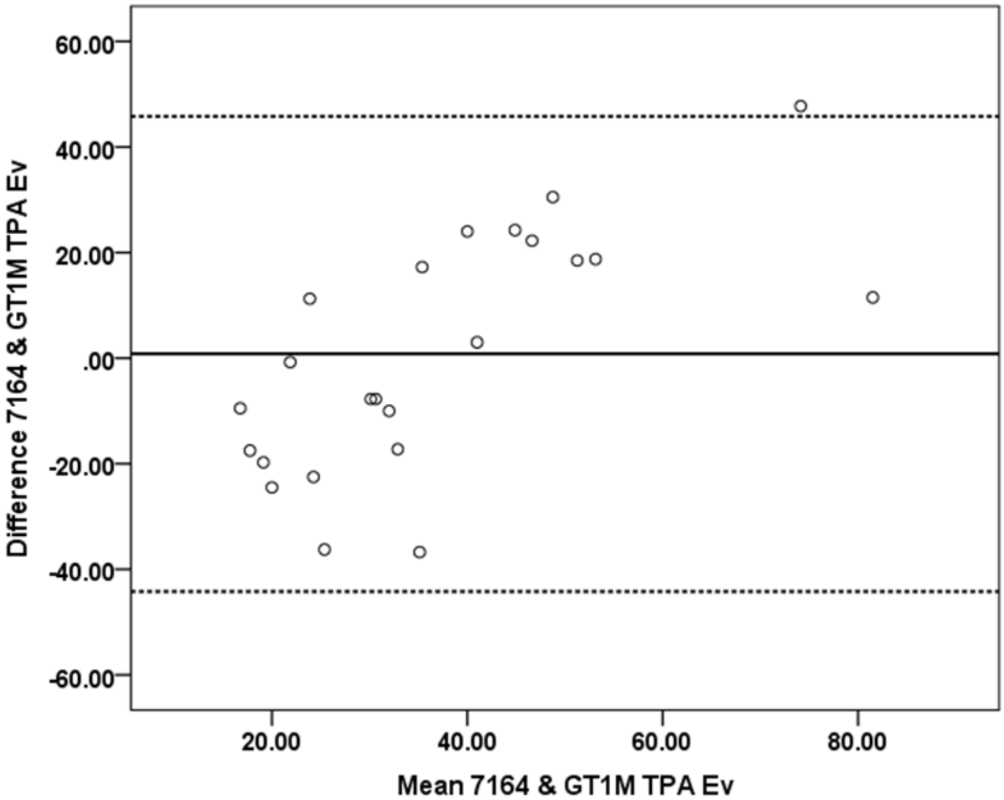
Table 9.6: Comparison between 7164 and GT1M accelerometers for time spent in sedentary and total physical activity with the Evenson et al. (2008) cut-points applied.

	Median (IQR) min		
Intensity	**7164**	**GT1M**	***dm* (LOA)**
Sed^{ev}	7.8 (8.8)	7.3 (7.5)*	-1.5 (-12.7 to 9.7)
TPA^{ev}	27.0 (44.8)	**34.5 (14)**	0.8 (-44.2 to 45.8)

*dm: mean difference; ev: Evenson et al. (2008); LOA: limits of agreement; Sed: Sedentary behaviour; TPA: Total physical activity. Values in bold not significantly different from 7164; * values significantly different from 7164 estimate.*

While there was no significant difference in the estimates of median time spent in total physical activity between the 7164 and the GT1M models, and the mean difference between measures was close to zero, there were, however, wide limits of agreement (-44.2 to 45.8 min) which are of concern. This is illustrated in a Bland and Altman plot presented in Figure 9.3. The data points above the line are where the 7164 recorded higher number of minute of total physical activity and those below the line are where the GT1M recorded a higher number of minutes of total physical activity.

Figure 9.3: Bland and Altman plot of minutes of total physical activity comparing 7164 accelerometer with GT1M with the Evenson et al. (2008) cut-points applied.



Ev: Evenson et al. (2008) cut-points applied; TPA: total physical activity. Solid lines depict mean difference in time and the dotted lines the limits of agreement (LOA) (mean difference (LOA): 0.8 (-44.to 45.8) min).

In conclusion, this study also supports the findings of Chapter 6 that the Puyau et al. (2002) cut-points for the GT1M data have the 'best' agreement with the CARS criterion measure of direct observation. However, if researchers are interested in the convergent validity of the GT1M with the 7164 accelerometer, the findings of this study suggest that applying a correction factor to the GT1M data (7164 = GT1M/0.91), as recommended by Corder et al. (2007), may be appropriate for acceptably accurate estimates of TPA. Applying the correction factor may go some way to resolving the issue of the output from the GT1M having lower cpm than the 7164 model, which, in this study, resulted in the GT1M recording more time as LPA and less time in MVPA in comparison to the 7164 accelerometer. These differences are largely dependent on the cut-point method applied. A summary of the key findings for Chapter 7 are presented in Table 9.7.

Table 9.7: Summary of key findings from Chapter 7.

Are different generations of Actigraph accelerometer comparable? (Chapter 7)	• Different generations of Actigraph (GT1M and 7164) are not comparable for estimates of sedentary behaviour, light physical activity or MVPA. • GT1M records more time as sedentary and less time as MVPA in comparison to the 7164 accelerometer. • There is good agreement between 7164 output and GT1M output for total physical activity once a correction factor is applied to the GT1M data (7164 = GT1M/0.91). • Applying a correction factor may in part resolve the issue of the output from the GT1M having lower cpm than the 7164 model.

9.6 WHAT IS THE RECOMMENDED WEAR TIME TO PROVIDE A RELIABLE ESTIMATE OF HABITUAL PHYSICAL ACTIVITY AND SEDENTARY BEHAVIOUR IN PRE-SCHOOL CHILDREN?

The aim of the final empirical chapter of this thesis was to determine the number of hours and days of monitoring required to provide a reliable estimate of habitual physical activity of pre-school children. The study also investigated whether inclusion of a weekend day was necessary for reliable estimates of habitual physical activity. Finally, the study aimed to investigate the influence that applying different criteria for non-wear time had on non eligibility rates from a sample and on estimates of total physical activity and sedentary behaviour. Analysis was conducted on accelerometry data collected over a 7-day period from 112 participants (mean (SD) age: 3.7 (0.2) y). Intra class correlation coefficients were calculated for a variety of numbers of hours (between 1 to 10 hours) and days (between 1 to 7 days) of data captured. Using the ICC single day values, a Spearman Brown Prophecy formula was used to predict the number of days needed to achieve an arbitrarily acceptable level of reliability of 0.7 (Hinkley et al. 2012b).

The results suggest that for the total sample there were no differences in accelerometer cpm between weekend and weekdays; however, some gender differences were apparent, such as girls being more active at weekends than during weekdays. During weekend days, children engaged in up to 1.2% more TPA than they did during weekdays. However, the difference in TPA was non significant ($p > 0.05$) and it is questionable whether it would be biologically meaningful.

Figure 8.4 (p 170) presented the wear time patterns of the young children and it was interesting to note that the wear time was similar for weekend and weekdays, with a similar gradient of curve for these days. This could suggest that younger children have similar wear time patterns during weekend and weekdays. This similarity in wear time pattern is in contrast to what has been observed in adolescents (Catellier et al. 2005), where the start times on the weekend days were much later, and the gradient of the slope of the graph for start times was more gradual, suggesting more variability of start time between participants. In the Catellier et al. (2005) study, 70% of sample participants were wearing their accelerometers by 12:00 noon on a weekend day, as opposed to 7.30 am for weekday. For the current study, 70% were wearing accelerometers by 9:00 am on a weekday and 10:00 am on a weekend day.

It was found that the inclusion of a weekend day had minimal influence on the reliability coefficient values and therefore the inclusion of a weekend day was deemed unnecessary for reliable estimates of habitual physical activity in this sample. While the ICC values were similar for all hours of data collection, 7 hours of data collection exhibited the highest ICC value (0.44) and using this wear time (7 hours), it was estimated that three days of data collection would achieve an ICC value of 0.70. Descriptive analysis of non-wear time revealed that the application of different thresholds of non-wear time resulted in varying estimates of total time spent physically active and sedentary. The influence that non-wear time criteria can have on apparent estimates of TPA and time spent in sedentary behaviour should be considered when comparing the outcomes of different studies.

Applying different criteria for non-wear time did not have a substantial impact on the number of excluded participants for 6 hour and 8 hour days when 4 or fewer days of data were collected. In these scenarios, only one or two participants were excluded when the different criteria for non-wear time were applied. However, when 10 hour days of data are collected, application of non-wear time criteria could have a considerable impact. For example, if data are collected over 3 days, applying the 10 minute non-wear time criterion (e.g. excluding 10 minutes of consecutive zeros in the data) would result in 35 participants being excluded from the final sample analysis, instead of 18 participants being excluded for a 60 minute non-wear time criterion (Chapter 8, Table 8.8).

The exclusion of different periods of non-wear time also had an influence on the estimates of time spent at different intensities. Excluding 10 minutes of non-wear time resulted in more time classified as TPA in comparison to when the 60 minute non-wear time criterion is applied. The opposite effect occurs with sedentary behaviour (e.g. excluding 60 minutes of non-wear results in more time classified as sedentary in comparison to when the 10 minute non-wear time criteria is applied). Currently only two studies, one by Esliger et al. (2005) and the other by Alhassan et al. (2008), have investigated the criteria for non-wear time in children and neither of these has been validated against a criterion method. However, both studies are in agreement in proposing 20 minutes as the criterion for non-wear time.

In conclusion, the results of the study presented in Chapter 8 suggest that there were no significant differences in daily average cpm between weekend and weekdays. Although some gender differences were apparent, such as boys having higher cpm than girls during weekdays ($p < 0.05$) and girls having significantly higher cpm during weekend days in comparison to weekdays (p = 0.03), there was no significant difference in percentage time per day spent in TPA between weekend and weekdays ($p > 0.05$). The pattern of wear time of accelerometers was similar for weekend and weekdays. One possible explanation for this is that younger children appear to be active for similar periods of time during weekend and weekdays and that this is different from what has been reported in studies of adolescents (Catellier et al. 2005). The inclusion of a weekend day had minimal influence on reliability coefficient values and therefore the inclusion of a weekend day is not considered necessary to reliably characterise the physical activity behaviour of this age-group. While the ICC values were similar for all hours of daily data collection, 7 hours of data collection had the highest single day ICC value (0.44) and it is estimated that 3 days of data collection would achieve an ICC value of 0.70. Descriptive analysis of the influence of applying different criteria for non-wear time revealed that the application of different thresholds of non-wear time did result in different estimates of average daily time spent in TPA and in sedentary behaviour. For example, using data from the same sample and excluding 10 minutes in comparison to excluding 60 minutes of consecutive zeros with a minimum of 10 hours of data collected would result in a mean difference (LOA) of 5.5 (-84.5 to 95.4) minutes in estimated TPA per day and mean difference (LOA) of 35.6 (-77.9 to 149.5) min/day being sedentary. The influence that applying different criteria for non-wear time has on estimates of TPA and time spent being sedentary has to be considered when comparing the findings of studies that have adopted different non-wear time criteria.

Table 9.8: Summary of key findings Chapter 8.

What is the recommended wear time to reliably estimate habitual physical activity and sedentary behaviour in pre-school children? (Chapter 8)	• Small non-significant difference between estimates of percentage time spent in total physical activity between weekend and weekdays. • Three days of data, for 7 hours per day, provide the minimum wear time criteria for a reliable estimate of habitual physical activity in pre-school children. • Inclusion of a weekend day in analysis is not necessary for a reliable estimate. • Different criteria for non-wear time had an impact on estimates of time spent in sedentary behaviour and total physical activity.

9.7 LIMITATIONS

The limitations of accelerometers have already been considered in the literature review chapter and this section will therefore focus on the limitations identified within the process of undertaking the studies within this thesis.

The first limitation to be considered is the small sample sizes used in the direct observation studies, in Chapters 4 to 6 (n = 31) and Chapter 7 (n = 23) which may influence the external validity of the findings and thus the generalisability from the sample observations to the wider population of pre-school aged children in Scotland. In addition, studies included in this thesis opportunistically recruited volunteers, and as such did not use a random sampling method. However, the mean values for height and weight were normally distributed and within the normal values expected for the pre-school population (ISD Scotland 2010). With an *a priori* sample size calculation, based on a standard study power calculation for power = 0.80, α level 0.05, an estimated 26 participants were deemed necessary to detect an effect size of 0.50 (Puyau et al. 2002). While the sample size in the Chapter 7 study was less than 26 (n = 23), preliminary analysis of 23 subjects revealed significant differences between models and therefore the study was felt to be sufficiently powered. Finally, small convenience samples are the norm in calibration and validation studies (Reilly et al. 2003; Sirard et al. 2005; Van Cauwenberghe et al. 2011), particularly those that involve direct observation, as this is recognised as being a time-consuming method (Oliver et al. 2007b).

One of the more recent developments in accelerometry technology is the improved memory storage capacity of current accelerometers, which means that it is now possible to collect data in 1-s epochs for extended periods of time. It remains unclear whether shorter epochs would offer greater precision in measurement of physical activity in pre-school children. A limitation of the current study is that the minimum epoch length used in the CARS criterion measure was at 15 s and therefore it was not possible to determine if shorter epochs would offer greater accuracy in measurement. More recent studies by Kahan et al. (2013) have used have used the Observation System for Recording Physical Activity in Children-Pre-school version (OSRAC-P; (Brown et al. 2006), in which data were coded over shorter epochs of 5-s. Although high inter-observer agreement of 90% and kappa means of 80%, suggesting substantial agreement, between physical activity level and the OSRAC-P observation system (Brown et al. 2006) are reported, this system does not involve continuous coding of the data. Instead the OSRAC-P incorporates a 5-s observation interval, followed by a 25-s recording interval, which may limit this approach as a criterion method. Other recent studies have made use of second-by-second coding of either the CARS (Oliver et al. 2009; Van Cauwenberghe et al. 2011), or an alternative direct observation scoring system (De Decker et al. 2013; Trost et al. 2012), both of which are loosely based on the CARS. Using a second-by-second approach to coding would offer an opportunity to use direct observation as a criterion at shorter epochs, against which cut-points could be validated. While this is promising for future studies which intend to use direct observation, the validity and reliability of these approaches has yet to be established, and the potential for observer error and the increased experimenter burden of second-by-second coding needs to be considered.

A further limitation of the current study is that the CARS scoring system does not differentiate between sitting and standing activities and both these stationary activities are classified as being ‘sedentary’ behaviour (Puhl et al. 1990). There is a growing argument that sitting and standing should be assessed separately, and in studies of adults the differences in energy expenditure between sitting and standing activities are argued to be important determinants of physical activity energy expenditure (Levine et al. 2008; Westerterp 2001). In some studies of young children standing is classified as a LPA (Trost et al. 2012) and in other studies ‘light’ intensity activities include sitting activities, for example sitting and writing (Pate et al. 2008). A recent consensus statement has argued that ‘true’ sedentary behaviour should be defined as sitting or reclining activities, where the resting metabolic rate is typically $\leq$ 1.5 METs (Sedentary Behaviour Research Network 2012; Tremblay et al. 2011).

De Decker et al. (2013) reported on the use of an observation scoring system which differentiated between sitting and standing still behaviours. These authors found that the ability to accurately differentiate between sitting and standing behaviours using the *activ*PAL™ (PAL Technologies, Glasgow, Scotland) and the Actigraph (GT1M) accelerometers was not conclusive. Given the lack of consensus in the literature, they recommended that further research is needed to establish whether or not stationary standing is, indeed, a ‘sedentary’ behaviour (De Decker et al. 2013).

The correct classification of sedentary behaviour has important implications for the classification of LPA. Differentiation between sitting and standing activities by using low threshold cut-points to classify sedentary behaviour which only includes sitting behaviours may be warranted, particularly given the lower cpm recorded by the GT1M in comparison to the 7164 accelerometer. There has been debate in the literature that sedentary behaviour is not dichotomous with physical activity and that they should be considered as separate constructs (Pate et al. 2011). Despite this, just as studies have been criticised for classifying all physical inactivity as 'sedentary time' (Sedentary Behaviour Research Network 2012), the low threshold for sedentary behaviour is being used to classify all 'other' behaviour above the threshold as part of 'total' physical activity.

The ability to distinguish between different sedentary activities warrants further investigation and one potential strategy is to make use of accelerometers such as the *activ*PAL™, which can record posture and activity (Davies et al. 2012a). The *activ*PAL™ has also been shown to produce similar estimates of physical activity and sedentary behaviour as the GT1M and GT3X Actigraph accelerometers on a group level in free-living data collected from pre-school children (n = 23, mean (SD) age: 4.5 (0.7) y) over 7 days (Martin et al. 2011). The *activ*PAL™ has also been found to have acceptable validity for identifying different postures in pre-school children (n = 30, mean (range) age: 4.2 (3.1 - 4.9) y) observed over 1 hour, being able to distinguish between lying, sitting, standing and walking activities (Davies et al. 2012a), as well as identifying the frequency of transitions between postures (Davies et al. 2012b). This is important, as evidence in adult studies suggests that the frequency of transitions between postures may have an influence on risk factors for cardiac disease (Healy et al. 2011; Tremblay et al. 2010).

A final limitation of the CARS is that the use of direct observation methods as a criterion measure has been questioned (Adolph et al. 2012; Oliver et al. 2007b). Adolph et al. (2012) argues that while direct observation is critical for identification of type of activity, this method lacks precision in quantifying intensity levels and thus energy expenditure. Moreover, Oliver et al. (2007b) argues that direct observation should be considered as a subjective method, as it relies on the observer to observe, interpret and code children's physical activity behaviour. Despite this, given the problems with interpreting energy expenditure in young children (see Chapter 1, p 2), observational methods offer a behavioural approach to calibration as an alternative. In addition, this approach is argued to be valuable in studies of young children, particularly if there is an interest in 'type' or 'patterns' of physical activity (Freedson et al. 2005). It has been argued that knowing the 'type' of activity that people are involved in is important as it may lead to more accurate estimates of energy expenditure where activity-specific predictions can be applied (Bassett et al. 2012).

Another limitation of the study is that the GT1M data were collected in 1-s epochs and then re-integrated into 5-, 15-, 30- and 60-s epochs. A recent publication by Kim et al. (2013) examined the effect of reintegration of smaller epochs (1-s epochs) into larger epochs (15-, 30-, and 60-s epochs) on activity counts and estimates of MVPA in accelerometry data collected from 31 pre-school children (3 to 5 y) over 1 school day. The authors suggest that there was very little difference in the group means for overall counts between smaller epochs reintegrated, and the larger epochs. One exception was with the estimates for MVPA using the Evenson et al. (2008) cut-points. Significant group mean differences in estimates of MVPA were found when 1-s epochs, reintegrated to 15-s epochs, were compared with data collected in 15-s epochs. However, the Cohen's effect sizes (*d*) were small (ranging from -0.2 to 0.1). While the effect sizes were small this could be a potential source of error when making comparisons between data collected in different epochs.

In the current study the Bland and Altman method (Bland and Altman 1986) was used to evaluate bias and limits of agreement between the CARS criterion measure and the accelerometry estimates of time spent in different intensities over the data collection period. Some calibration studies have also made use of analysis using receiver operating characteristics (ROC) curve analysis to determine the sensitivity (true positive rate) and specificity (false positive rate) of the accelerometry output against the CARS (De Decker et al. 2013; Evenson et al. 2008; Kahan et al. 2013; Reilly et al. 2003). This method was not adopted in the current study due to concerns with synchronisation of the GT1M output which was identified during the mechanical calibration processes (Appendix Table II.xi and Appendix Table II.xii p. 307). During the mechanical calibration of the GT1M accelerometers, it was noted that the GT1M accelerometers were not accurately synchronising with the PC time, with up to 15 s discrepancy between that which was recorded on the GT1M and the PC time. This became apparent as the mechanical oscillations were timed with the PC and in simultaneous calibration undertaken with the 7164 and the RT3 models where no similar problems of synchronisation were encountered. As the CARS was coded using video data which was also synchronised with the PC time, this questions the accuracy of each 15-s period of accelerometer output synchronising with each 15-s period of video data output. For this reason it was not appropriate to use the ROC analysis approach, as there may be errors in the synchronisation between the two measurements (accelerometry output and video data output), which would make classification accuracy doubtful.

9.8 SUGGESTIONS FOR FUTURE STUDIES

In this section the areas for future study will be considered making reference to the advances in accelerometry technology that have arisen since publication of the chapters within this thesis.

The current thesis predominantly made use of the GT1M Actigraph accelerometer model, which has been superseded by the production of the triaxial GT3X and GT3X+ Actigraph accelerometers. Given that the GT3X and the GT3X+ are now available, it is likely that future studies may focus on the validity and reliability of these instruments to accurately quantify physical activity behaviour of pre-school children.

Since 2010, the GT1M and GT3X have been able to provide output as raw acceleration data so that the data represents the *G*-force sampled every 0.033s in the pre-filtered raw mode. The bandwidth filter has been found to influence accelerometry measurements and during mechanical testing using an orbital shaker, a plateau effect has been observed at higher shaker frequencies and suppression of output has been seen during lower shaker frequencies (Chen et al. 2012; Rothney et al. 2008). Using accelerometers in an unfiltered mode may be important for accurate detection of low and high intensity activity, where the filter can reduce sensitivity of measurement (Chen et al. 2012).

Another recent development is the availability of a low frequency extension option for the GT1M and GT3X accelerometers to improve their sensitivity at lower frequencies. Early studies with adults suggest that this may improve comparability between the 7164 and the newer generations of Actigraph, however, this has to be investigated with different populations, including pre-school children (Ried-Larsen et al. 2012).

With the availability of raw accelerometry data it has been argued that rather than reporting the 'count' output which, as discussed, varies between the models of accelerometer and trying to convert this to a biologically meaningful format, future accelerometers should provide data in standardised units such as gravitational constant (G, $m \cdot s^{-2}$) or time-integrated units ($m \cdot s^{-1}$). This would allow for greater ease of comparison between accelerometry output between models. The use of raw accelerometry data instead of activity counts had been proposed for incorporation within a consensus statement at the 2009 'Objective Measurement of Physical Activity: Best Practice and Future Directions' conference (John and Freedson 2012). One possible concern with using raw acceleration data is how to manage the raw acceleration data and convert this into a meaningful format for the ‘end user’ be it researcher or policy-maker.

Advances have, however, also been made in accelerometer data processing, with the development of more sophisticated approaches to data modelling analysis. One approach is ‘pattern recognition’, which makes use of raw acceleration data or 1-s epochs of data to identify types of behaviour (Bonomi et al. 2009; Pober et al. 2006; Staudenmayer et al. 2009; Zhang et al. 2003) and energy expenditure (Bonomi et al. 2009; Rothney et al. 2007; Staudenmayer et al. 2009; Zhang et al. 2004). These pattern recognition approaches to processing accelerometry data have been argued to have better accuracy at estimating energy expenditure than other methods, such as single-regression or double-regression equations (Bassett et al. 2012). Promising results have been found with Artificial Neural Network (ANN) a pattern recognition approach to data processing in a study of children (n = 58, mean (SD) age=11.0 (0.7) y) (de Vries et al. 2011). The authors reported 76.8% accuracy of different activities from a hip worn triaxial accelerometer (GT3X) (de Vries et al. 2011). This area warrants further investigation in studies of pre-school children to determine if this will offer an accurate means of classifying physical activity behaviour.

The latest monitor from Actigraph the GT3X+ has a greater storage capacity than the earlier models and is waterproof, which means that the child can continue to wear the accelerometer for water based activities. It also incorporates a light sensor so that sleeping and awake times can be estimated. These factors could allow for 24 hour monitoring over extended periods of time. The GT3X+ monitor collects data in a raw unfiltered mode and gives the researcher the ability to select one of 10 different sampling rates from 10 to 100 Hz (John and Freedson 2012). Increasing the available options is useful for researchers, however, agreement on the optimal sampling rate to use will need to be determined. In addition, while this is advancing the field of accelerometry, it also introduces more options for researchers for which decisions need to be made and agreed upon and the need for a consensus on protocols becomes even more imperative (Cain et al. 2013).

Finally, there is recent debate about whether wrist worn accelerometers, instead of hip worn monitors may improve compliance (Routen et al. 2012), and account for energy expenditure from upper limb movements. While the Actiwatch has been available for a number of years, the GT3X+ is now available as a wrist worn monitor. The Actiwatch has been used successfully over extended periods of time with school-aged children (Nyberg et al. 2009), however, the output between wrist and hip worn accelerometers differs and is not comparable (Routen et al. 2012). Research is therefore needed to calibrate wrist worn accelerometers such as the GT3X+ and to determine the feasibility of their use over extended periods of time with pre-school children.

In the final study chapter, the application of different criteria for non-wear time on estimates of sedentary behaviour and total physical activity was highlighted. What this study did not determine is which threshold for non-wear time is most accurate. Given the premise that pre-school children are unlikely to be stationary for more than 10 minutes has not been established, future studies should investigate which non-wear time threshold is appropriate to use with these children.

Advances in accelerometry technology to include the incorporation of contextual information on the location of activity by Global Positioning Systems are being developed and it is argued that this type of information integrated with information on the purpose of an activity will provide a useful adjunct to surveillance data (Matthews et al. 2012). There is also an emergence of multiple sensor systems which provide the ability to simultaneously collect physiological variables such as body heat and heart-rate with accelerometry. The additional benefits of these new advances in the measurement of physical activity are under investigation (Matthews et al. 2012).

Despite the availability of the GT3X and GT3X+, several on-going, large scale longitudinal accelerometry studies are using and continue to use the 7164 and GT1M Actigraph accelerometers such as the Avon Longitudinal Study of Parents and Children (ALSPAC) (Riddoch et al. 2007), the Millennium Cohort Study in the UK (Basterfield et al. 2011) and the European Youth Heart Study (EYHS) in four European Countries: Estonia, Denmark, Portugal and Norway (Brage et al. 2004b; Moller et al. 2009; Riddoch et al. 2004). Establishing consensus on how these accelerometers are being utilised is therefore essential to allow for meaningful longitudinal analysis of physical activity and sedentary behaviour of populations. Future studies will, of course, need to investigate the comparability between the GT3X and GT3X+ with the 7164 and GT1M, particularly if raw accelerometry data are to be used. In addition, how this data will be translated into a format which gives insight into physical activity and sedentary behaviour in populations needs to be determined. In deciding which measurement approach to adopt, researchers need to weigh up the relative merits of measurement accuracy through deploying high cost hardware and software, with high participant burden (through use of multi sensor systems, for example), against the expediency of the lower-cost, lower participant burden but possibly less precise measurement accuracy associated with simpler methods (such as single sensor accelerometry) (Matthews et al. 2012; Trost et al. 2005).

A summary of the limitations of the thesis are given in Table 9.9 and suggestions for future research are given in Table 9.10.

Table 9.9: Summary of limitations of thesis.

Limitations
• Small sample sizes in direct observation studies may limit the scope for generalisation.
• Use of coding CARS at 15-s epochs. Unable to determine accuracy of epochs less than 15 s.
• Limitations with the CARS classification of sedentary behaviour, in particular differentiating between sitting and standing activities.
• Objectivity of direct observation systems to accurately estimate intensity level has been questioned.
• Availability of GT3X & GT3X+ with new measurement possibilities becoming available.
• Problems with synchronisation of GT1M accelerometers meant ROC analysis was not appropriate.
• Possible concerns about comparability of data reintegrated from 1-s epochs with data collected in 15-s epochs.
• Accurate identification of non-wear time and differentiating this from sedentary time has not been determined.

Table 9.10: Summary of suggestions for future research

Future studies
• Need for future studies to consider coding the CARS at shorter epochs e.g. 1s and 5s as the criterion method to determine if shorter epochs offer greater accuracy. • Need for validation of a direct observation scoring system which can differentiate between sitting and standing activities. • Future research needed to validate thresholds for the GT3X and GT3X+. • Investigate the use of low frequency extension for the GT1M and GT3X to improve comparability between generations of Actigraph monitors during free-living activities. • Validation studies of the raw unfiltered accelerometry data against criterion measures of physical activity are needed. • Further research is needed to identify sitting and standing activities, transition activities through use of accelerometers which incorporate an inclinometer such as the *activ*PAL™ or data processing using modelling techniques such as Artificial Neural Network to improve accuracy of measurement of physical activity and sedentary behaviour of pre-school children. • Studies to determine the 'optimal' non-wear time criteria for pre-school children are necessary.

9.9 RECOMMENDATIONS FOR FUTURE PRACTICE

The findings of this thesis contribute to the existing knowledge relating to standardisation protocols for accelerometry use in free-living habitual physical activity of pre-school children. The recommendations arising from this research suggest that researchers should use short epochs of at least 15 s, as the current research suggests these offer a more accurate means of measuring physical activity intensity in pre-school children. Using 60-s epochs appears to underestimate true levels of MVPA. While researchers may want to consider using epochs which are shorter than 15 seconds (e.g. 1- and 5-s epochs), they should be used with cut-points which have been validated for shorter epochs. To date the shortest epochs used to validate cut-points have been at 15 s. Policy-makers and researchers need to be aware of the implications of studies that have made use of different epochs, as this may have an impact on the estimates of physical activity, and influence comparability between studies.

For research purposes, the evidence in this thesis suggests that uniaxial accelerometers are sufficiently accurate in quantifying physical activity behaviour in pre-school children. If researchers are using the GT3X and GT3X+, they appear to be comparable with the GT1M if used on a uniaxial setting (Kaminsky and Ozemek 2012; Robusto and Trost 2012; Vanhelst et al. 2012). This means that the cut-points validated for the GT1M could be used with the GT3X and GT3X+ in a uniaxial mode as these monitors have the same internal technology, same specifications including the band-pass filter, analogue-to-digital converter (John and Freedson 2012) and have been found to respond similarly in free-living settings (Kaminsky and Ozemek 2012; Ried-Larsen et al. 2012; Vanhelst et al. 2012). If researchers intend to use triaxial accelerometers in their vector magnitude mode, model-specific cut-points should be used, noting that validated cut-points are not yet available for the GT3X and GT3X+. If researchers and policy-makers are interested in estimates of time spent in different intensities then the Puyau et al. (2002) cut-points provide the most accurate estimates at a group level for MVPA. However, while the Puyau et al. (2002) cut-points were accurate for sedentary and LPA as well as TPA, it is possible that these overestimate time spent in sedentary behaviour (and subsequently underestimate LPA). The current cut-points were calibrated against the CARS which includes both stationary standing and sitting behaviours within the 'sedentary' category.

Researchers should also be aware that the output from the GT1M and the 7164 are not comparable and these monitors cannot be assumed to provide the same estimates of time spent at different intensities. This means that the cut-points validated for the 7164 are not necessarily valid for the GT1M and ideally cut-points should be validated for the GT1M accelerometer. Application of a correction factor did not resolve this issue for estimates of MVPA, but may potentially allow for comparison of TPA between studies if the correction factor suggested by Corder et al. (2007) (7164 = GT1M/0.91) is applied to the GT1M data. While the availability of low frequency extension for the GT1M accelerometer may improve comparability with the 7164 accelerometer, this has not been verified for free-living with pre-school children. Researchers and policy-makers should interpret the output from studies which have used the different models of accelerometer, or use different models within the same study cautiously. This has important implications, particularly for accuracy of physical activity monitoring within large scale longitudinal studies.

When conducting physical activity research, researchers commonly encounter problems with compliance of subjects with their data collection protocols. As a result researchers have to make a series of decisions around what constitutes the minimum 'wear time' for inclusion of a subject's data within their analysis. These decisions are as follows: i) how many hours constitute a 'valid' day, ii) how may days of data collection are necessary to provide a reliable estimate of typical habitual physical activity of the population, iii) what definitions of non-wear time are applied and how this 'missing' data is dealt with. Researchers and policy makers need to be aware that the wear time decisions can affect the study sample size and estimates of physical activity and sedentary behaviour (Mâsse et al. 2005). This can mean that the outcomes of different studies are not necessarily comparable if different decisions were made; this has implications for longitudinal surveillance of physical activity and sedentary behaviour within populations.

Results suggest that, for the sample of pre-school children analysed in this thesis, 7 hours of data collection constitute a 'valid' day. Three days of data collection for 7 hours per day provide adequate reliability of data (r = 0.70). Inclusion of a weekend day was found not to be necessary for a reliable estimate of habitual activity in pre-school children. This means that subjects' data were included in the final analysis even if they did not include a weekend day. This decision influences the sample size as, out of 104 subjects with 7 hours of data for 3 days, only 80 subjects had one weekend day of data. Excluding participants without a weekend day would mean that the sample was reduced by 29% from the original sample size (n = 112). Researchers and policy makers should be aware that more stringent criteria for wear time may have implications for sample sizes included and on subsequent estimates of physical activity.

The criteria for non-wear time should be consistent across studies and current evidence from the literature suggests that 20 minutes of consecutive zeros should be adopted (Alhassan et al. 2008; Esliger et al. 2005). Researchers and policy-makers need to be aware that the exclusion of different periods of non-wear time also influences sample size as well as inflating or deflating estimates of MVPA and sedentary behaviour. A summary of the recommendations are presented in Table 9.11.

Table 9.11: Recommendations arising from thesis for researchers and policy-makers.

Recommendations
1. Epoch: 15-s epochs should be used until further evidence establishes whether shorter epochs are more accurate.
2. Uniaxial or triaxial: Uniaxial accelerometers are sufficiently accurate for measurement of PA and sedentary behaviour with pre-school children.
3. Cut-points: Puyau et al. (2002) cut-points provided the 'best' agreement with CARS criterion measure and should be adopted if researchers/policy-makers are interested in the patterns and types of PA.
4. Different generations of accelerometer: Output from the 7164 and GT1M are not comparable and the cut-points from the 7164 are not necessarily validated for use with the GT1M. Application of a correction factor to the GT1M data (GT1M/0.91) may allow for comparison of total physical activity between different generations of models.
5. Wear time: Three days of 7 hours of data collection are needed to reliably estimate total physical activity.
6. Inclusion of weekend day: The inclusion of a weekend day in the analysis is not necessary for reliable estimates of habitual physical activity.
7. Non-wear time: Criteria for non-wear time of 20 minutes should be adopted until further evidence establishes the best criteria to use with pre-school children.
8. Researchers and policy-makers need to be aware of the influence that decisions relating to points 1 to 7 above will have on output and carefully examine the approaches adopted in studies.
9. Researchers need to be explicit in their dissemination of findings about the methodological decision related to points 1 to 7.
10. Where possible, consistency within longitudinal studies should be adopted, such as use of the same epochs, same cut-points.
11. The research community needs to move towards establishing standardised protocols for data collection and processing.

CHAPTER 10 : CONCLUSIONS

The ability to accurately measure free-living physical activity and sedentary behaviour is crucial for any investigation in which physical activity is observed or is an intervention or exposure variable of interest (Strath et al. 2012). Accelerometers are one means of objectively measuring physical activity and sedentary behaviour. Accelerometers can characterise the intensity, duration and frequency of physical activity and sedentary behaviour and can be used across large populations in a reasonably cost-effective, non-invasive way with low participant or researcher burden involved. They are therefore ideal for assessment of physical activity and sedentary behaviour in population-wide studies and allow for the collection of essential statistics for public health planning and intervention (Strath et al. 2012). However, several outstanding questions around the use of accelerometers with pre-school children exist (Cliff et al. 2009b). This thesis has made progress in addressing six key methodological questions: are shorter epochs more accurate for use with young children (Chapter 3)? Which epochs are most accurate for use with pre-school children (Chapter 4)? Are there advantages of using triaxial over uniaxial accelerometers for pre-school children (Chapter 5)? Which cut-points are most accurate for use with pre-school children (Chapter 6)? Are different generations of Actigraph accelerometers comparable (Chapter 7)? What is the recommended wear time to provide a reliable estimate of habitual physical activity and sedentary behaviour in pre-school children (Chapter 8)?

Recommendations based on the results of the individual studies have been made for researchers and for policy-makers. Fundamental to the recommendations however, is the need for researchers to clearly document their approaches to data collection and processing so that the influence of the decisions they have made can be considered by the reader. Ideally, the research community needs to move towards a consensus around measurement to allow for standardisation across studies, in particular around which cut-points to apply. It is only by adopting standardised approaches to accelerometry assessment that comparison between outcomes of studies will be meaningful. Despite this, accelerometers offer the opportunity to objectively gather data on physical activity and sedentary behaviour. This method is important for surveillance of physical activity and sedentary behaviours in populations over time, and has added to the knowledge and understanding of the associations of physical activity and sedentary behaviour with health, which was not previously available through self-report methods.

Finally, consensus over what behaviours contribute to 'light' intensity physical activity is particularly important for pre-school children given the recommendations for health which relate to total physical activity for which LPA and MVPA are combined (Australian Government, Department of Health and Ageing 2009; Canadian Society of Exercise Physiology 2012; Department of Health, Physical Activity, Health Improvement and Protection 2011). Earlier debate in the literature regarding reaching agreement on what MVPA is (Guinhouya et al. 2006; Guinhouya and Hubert 2008) should now be focusing on what light intensity physical activity constitutes in young children, as this provides the threshold used to determine total physical activity. Until this is resolved, inconsistencies in apparent outcomes from assessment of total physical activity in populations of pre-school children will continue.

REFERENCES

Aaron, D.J., Kriska, A.M., Dearwater, S.R., Cauley, J.A., Metz, K.F. and LaPorte, R.E. 1995. Reproducibility and validity of an epidemiologic questionnaire to assess past year physical activity in adolescents. *American Journal of Epidemiology,* 142 (2) pp.191-201.

Adolph, A.L., Puyau, M.R., Vohra, F.A., Nicklas, T.A., Zakeri, I.F. and Butte, N.F. 2012. Validation of uniaxial and triaxial accelerometers for the assessment of physical activity in preschool children. *Journal of Physical Activity and Health,* 9 (7) pp.944-953.

Ainsworth, B.E., Haskell, W.L., Whitt, M.C., Irwin, M.L., Swartz, A.M., Strath, S.J., O'Brien, W.L., Bassett, D.R.,Jr., Schmitz, K.H., Emplaincourt, P.O., Jacobs, D.R.,Jr. and Leon, A.S. 2000. Compendium of physical activities: an update of activity codes and MET intensities. *Medicine and Science in Sports and Exercise,* 32 (Suppl 9) pp.S498-S504.

Alhassan, S., Sirard, J.R. and Robinson, T.N. 2007. The effects of increasing outdoor play time on physical activity in Latino preschool children. *International Journal of Pediatric Obesity,* 2 (3) pp.153-158.

Alhassan, S., Sirard, J.R., Spencer, T.R., Varady, A. and Robinson, T.N. 2008. Estimating physical activity from incomplete accelerometer data in field studies. *Journal of Physical Activity and Health,* 5 (Suppl 1) pp.S112-S125.

American Academy of Pediatrics. 2001. Children, adolescents and television. *Pediatrics,* 107 (2) pp.423-426.

American Academy of Pediatrics. 2006. Active healthy living: prevention of childhood obesity through increased physical activity. *Pediatrics,* 117 (5) pp.1834-1842.

Analog Devices, I. 2007. *ADXL320, Small and Thin ± 5 g Accelerometer iMEMS® Accelerometer.* [Online] Analog Devices, Inc. Available at: http://nees.unr.edu/Documents/nees/instrumentation/acceleration/adxl320.pdf [Accessed July 3 2007].

Andersen, L.B., Harro, M., Sardinha, L.B., Froberg, K., Ekelund, U., Brage, S. and Anderssen, A. 2006. Physical activity and clustered cardiovascular risk in children: a cross-sectional study (The European Youth Heart Study). *The Lancet,* 368 pp.299-304.

Anderson, C.B., Hagströmer, M. and Yngve, A. 2005. Validation of the PDPAR as an adolescent diary: the effect of accelerometer cut-points. *Medicine and Science in Sports and Exercise,* 37 (3) pp.1224-1230.

Atkinson, G. and Nevill, A.M. 1998. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine*, 26 (4) pp.217-238.

Australian Government, Department of Health and Ageing., 2009. *Get up and grow: healthy eating and physical activity for early childhood. Directors/coordinators book.* [online].Canberra: Australian Government, Department of Health and Ageing [Accessed January 16 2013]. Available from: http://www.health.gov.au/internet/main/publishing.nsf/Content/phd-gug-directorscoord.

Ayabe, M., Kumahara, H., Morimura, K. and Tanaka, H. 2013. Epoch length and physical activity bout analysis: an accelerometry research issue. *BMC Research Notes* 6 (20), Available at: http://www.biomedcentral.com/1756-0500/6/20 [Accessed April 8 2013].

Bailey, R.C., Olson, J., Pepper, S.L., Porszasz, J., Barstow, T.J. and Cooper, D.M. 1995. The level and tempo of children's physical activities: an observational study. *Medicine and Science in Sports and Exercise,* 27 (7) pp.1033-1041.

Baquet, G., Stratton, G., Van Praagh, E. and Berthoin, S. 2007. Improving physical activity assessment in prepubertal children with high-frequency accelerometry monitoring: a methodological issue. *Preventive Medicine,* 44 (2) pp.143-147.

Baranowski, T., Bouchard, C., Bar-Or, O., Bricker, T., Heath, G., Kimm, S.Y., Malina, R., Obarzaned, E., Pate, R. and Strong, W.B. 1992. Assessment, prevalence, and cardiovascular benefits of physical activity and fitness in youth. *Medicine and Science in Sports and Exercise,* 24 (Suppl 6) pp.S237-S247.

Baranowski, T. and de Moor, C. 2000. How many days was that? Intra-individual variability and physical activity assessment. *Research Quarterly for Exercise and Sport,* 71 (74-8) pp.S74-S78.

Baranowski, T., Thompson, W.O., DuRant, R.H., Baranowski, J. and Puhl, J. 1993. Observations on physical activity in physical locations: age, gender, ethnicity, and month effects. *Research Quarterly for Exercise and Sport,* 64 (2) pp.127-133.

Bar-Or, O. and Malina, R.M. 1995. Activity, fitness and health of children and adolescents. In: Cheung, L.W.Y. and Richmond, J.B. eds. *Child health, nutrition and physical activity*. Champaign, IL: Human Kinetics, pp. 79-112.

Bascardí-Gascón, M., Reveles-Rojas, C., López Gail, W., Crawford, P. and Jiménez-Cruz, A. 2011. Validity of a physical activity questionnaire used with parents of preschool children in Mexico. *Nutrición Hospitalaria,* 26 (1) pp.244-245.

Bassett, D.R.,Jr., Ainsworth, B.E., Swartz, A.M., Strath, S.J., O'Brien, W.L. and King, G.A. 2000. Validity of four motion sensors in measuring moderate intensity physical activity. *Medicine and Science in Sports and Exercise,* 32 (Suppl 9) pp.S471-S480.

Bassett, D.R.,Jr., Rowlands, A. and Trost, S.G. 2012. Calibration and validation of wearable monitors. *Medicine and Science in Sports and Exercise,* 44 (Suppl 1) pp.S32-S38.

Basterfield, L., Adamson, A.J., Frary, J.K., Parkinson, K.N., Pearce, M.S., Reilly, J.J. and Gateshead Millennium Study Core Team. 2011. Longitudinal study of physical activity and sedentary behavior in children. *Pediatrics,* 127 (1) pp.e24-e30.

Basterfield, L., Pearce, M.S., Adamson, A.J., Frary, J.K., Parkinson, K.N., Wright, C.M. and Reilly, J.J. 2012. Physical activity, sedentary behavior, and adiposity in English children. *American Journal of Preventive Medicine,* 42 (5) pp.445-451.

Batterham, A.M. and George, K.P. 2000. Reliability in evidence-based clinical practice: a primer for allied health professions. *Physical Therapy in Sports,* 1 (2) pp.54-61.

Beets, M.W., Bornstein, D., Dowda, M. and Pate, R.R. 2011a. Compliance with national guidelines for physical activity in U.S. preschoolers: measurement and interpretation. *Pediatrics,* 127 (4) pp.658-664.

Beets, M.W., Morgan, C.F., Banda, J.A., Bornstein, D., Byun, W., Mitchell, J., Munselle, L., Rooney, L., Beighle, A. and Erwin, H. 2011b. Convergent validity of pedometer and accelerometer estimates of moderate-to-vigorous physical activity of youth. *Journal of Physical Activity and Health,* 8 (Suppl 2) pp.295-S305.

Beets, M.W., Patton, M.M. and Edwards, S. 2005. The accuracy of pedometer steps and time during walking in children. *Medicine and Science in Sports and Exercise,* 37 (3) pp.513-520.

Bender, J.M., Brownson, R.C., Elliott, M.B. and Haire-Joshu, D.L. 2005. Children's physical activity: using accelerometers to validate a parent proxy record. *Medicine and Science in Sports and Exercise,* 37 (8) pp.1409-1413.

Benham-Deal, T. 2005. Preschool children's accumulated and sustained physical activity. *Perceptual & Motor Skills,* 100 (2) pp.443-443.

Biddle, S., Cavill, N. and Sallis, J. 1998. Policy framework for young people and health-enhancing physical activity. In: Biddle, S., Sallis, J. and Cavill, N. eds. *Young and active? Young people and health-enhancing physical activity: evidence and implications.* London, UK: Health Education Authority, pp. 3-16.

Biddle, S.J., Pearson, N., Ross, G.M. and Braithwaite, R. 2010. Tracking of sedentary behaviours of young people: a systematic review. *Preventive Medicine,* 51 (5) pp.345-351.

Bland, J.M. and Altman, D.G. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet,* 1 (8476) pp.307-310.

Bonomi, A.G., Goris, A.H., Yin, B. and Westerterp, K.R. 2009. Detection of type, duration, and intensity of physical activity using an accelerometer. *Medicine and Science in Sports and Exercise,* 41 (9) pp.1770-1777.

Bornstein, D.B., Beets, M.W., Byun, W. and McIver, K. 2011. Accelerometer-derived physical activity levels of preschoolers: a meta-analysis. *Journal of Science and Medicine in Sport,* 14 (6) pp.504-511.

Bower, J.K., Hales, D.P., Tate, D.F., Rubin, D.A., Benjamin, S.E. and Ward, D.S. 2008. The childcare environment and children's physical activity. *American Journal of Preventive Medicine,* 34 (1) pp.23-29.

Brage, S., Brage, N., Wedderkopp, N. and Froberg, K. 2003. Reliability and validity of the Computer Science and applications accelerometer in a mechanical setting. *Measurement in Physical Education and Exercise Science,* 7 (2) pp.101-119.

Brage, S., Wedderkopp, N., Ekelund, U., Franks, P.W., Wareham, N.J., Andersen, L.B. and Froberg, K. 2004a. Features of the metabolic syndrome are associated with objectively measured physical activity and fitness in Danish children: The European Youth Heart Study (EYHS). *Diabetes Care,* 27 (9) pp.2141-2148.

Brage, S., Wedderkopp, N., Ekelund, U., Franks, P.W., Wareham, N.J., Andersen, L.B. and Froberg, K. 2004b. Objectively measured physical activity correlates with indices of insulin resistance in Danish children. The European Youth Heart Study (EYHS). *International Journal of Obesity*, 28 (11) pp.1503-1508.

Brage, S., Wedderkopp, N., Franks, P.W., Andersen, L.B. and Froberg, K. 2003. Reexamination of validity and reliability of the CSA monitor in walking and running. *Medicine and Science in Sports and Exercise*, 35 (8) pp.1447-1454.

Bringolf-Isler, B., Mäder, U., Ruch, N., Kriemler, S., Grize, L. and Braun-Fahrländer, C. 2012. Measuring and validating physical activity and sedentary behavior comparing a parental questionnaire to accelerometer data and diaries. *Pediatric Exercise Science*, 24 (2) pp.229-245.

Brown, H.B., Pfeiffer, K.A., McIver, K.L., Dowda, M., Almeida, J.M. and Pate, R.R. 2006. Assessing children's physical activity: the Observational System for Recording Physical Activity in children-preschool version. *Research Quarterly for Exercise and Sport*, 77 (2) pp.167-176.

Burdette, H.L., Whitaker, R.C. and Daniels, S.R. 2004. Parental report of outdoor playtime as a measure of physical activity in preschool-aged children. *Archives of Pediatrics and Adolescent Medicine*, 158 (4) pp.353-357.

Burdette, H.L. and Whitaker, R.C. 2005. Resurrecting free play in young children: looking beyond fitness and fatness to attention, affiliation, and affect. *Archives of Pediatrics and Adolescent Medicine*, 159 (1) pp.46-50.

Bürgi, F., Meyer, U., Niederer, I., Ebenegger, V., Marques-Vidal, P., Granacher, U., Kriemler, S. and Kriemler, S. 2010. Socio-cultural determinants of adiposity and physical activity in preschool children: a cross-sectional study. *BMC Public Health* [online] 10 pp.733. Available at: http://www.biomedcentral.com/content/pdf/1471-2458-10-733.pdf [Accessed July 25 2013].

Butte, N.F., Puyau, M.R., Adolph, A.L., Vohra, F.A. and Zakeri, I. 2007. Physical activity in nonoverweight and overweight Hispanic children and adolescents. *Medicine and Science in Sports and Exercise*, 39 (8) pp.1257-1266.

Cain, K.L., Sallis, J.F., Conway, T.L., Van Dyck, D. and Calhoon, L. 2013. Using accelerometers in youth physical activity studies: a review of methods. *Journal of Physical Activity and Health*, 10 (3) pp.437-450.

Canadian Society of Exercise Physiology., 2012. *Canadian physical activity guidelines: for the early years - 0-4 years.* [online].Ottawa: Canadian Society for Exercise Physiology [Accessed January 16 2013]. Available from: http://www.csep.ca/CMFiles/Guidelines/CSEP-InfoSheets-early-years-ENG.pdf.

Cardon, G.M. and Bourdeaudhuij, I.M. 2008. Are preschool children active enough? Objectively measured physical activity levels. *Research Quarterly for Exercise and Sport*, 79 (3) pp.326-332.

Cardon, G. and De Bourdeaudhuij, L. 2007. Comparison of pedometer and accelerometer measures of physical activity in preschool children. *Pediatric Exercise Science,* 19 (2) pp.205-214.

Carson, V. and Spence, J.C. 2010. Seasonal variation in physical activity among children and adolescents: a review. *Pediatric Exercise Science,* 22 (1) pp.81-92.

Caspersen, C.J., Powell, K.E. and Christenson, G.M. 1985. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Reports,* 100 (2) pp.126-131.

Catellier, D.J., Hannan, P.J., Murray, D.M., Addy, C.L., Conway, T.L., Yang, S. and Rice, J.C. 2005. Imputation of missing data when measuring physical activity by accelerometry. *Medicine and Science in Sports and Exercise,* 37 (11) pp.S555-S562.

Cavill, N., Kahlmeier, S. and Racioppi, F. eds. 2006. *Physical activity and health in Europe evidence for action.* Denmark: World Health Organisation.

Chen, K.Y. and Bassett, D.R.,Jr. 2005. The technology of accelerometry-based activity monitors: current and future. *Medicine and Science in Sports and Exercise,* 37 (Suppl11) pp.S490-S500.

Chen, K.Y., Janz, K.F., Zhu, W. and Brychta, R.J. 2012. Redefining the roles of sensors in objective physical activity monitoring. *Medicine and Science in Sports and Exercise,* 44 (1 Suppl 1) pp.S13-S23.

Chen, X., Sekine, M., Hamanishi, S., Wang, H., Hayashikawa, Y., Yamagami, T. and Kagamimori, S. 2002. The validity of nursery teachers' report on the physical activity of young children. *Journal of Epidemiology,* 12 (5) pp.367-374.

Choi, L., Lui, Z., Matthews, C.E. and Buchowski, M.S. 2011. Validation of accelerometer wear and nonwear time classification algorithm. *Medicine and Science in Sports and Exercise,* 43 (2) pp.357-364.

Chu, E.Y., McManus, A.M. and Yu, C.C. 2007. Calibration of the RT3 accelerometer for ambulation and nonambulation in children. *Medicine and Science in Sports and Exercise,* 39 (11) pp.2085-2091.

Cliff, D.P. and Okely, A.D. 2007. Comparison of two sets of accelerometer cut-off points for calculating moderate-to-vigorous physical activity in young children. *Journal of Physical Activity and Health,* 4 (4) pp.509-513.

Cliff, D.P., Okely, A.D., Smith, L.M. and McKeen, K. 2009a. Relationship between fundamental movement skills and objectively measured physical activity. *Pediatric Exercise Science,* 21 (4) pp.436-449.

Cliff, D.P., Reilly, J.J. and Okely, A.D. 2009b. Methodological considerations in using accelerometers to assess habitual physical activity in children aged 0-5 years. *Journal of Science and Medicine in Sport,* 12 (5) pp.557-567.

Cliff, D., P., Okely, A., D., Morgan, P., J., Steele, J., R., Jones, R., A., Colyvas, K. and Baur, L., A. 2011. Movement skills and physical activity in obese children: randomized controlled trial. *Medicine and Science in Sports and Exercise,* 43 (1) pp.90-100.

Cohen, J. 1988. *Statistical power analysis for behavioural sciences.* 2nd ed. New York: Academic Press.

Cohen, J. 1992. A power primer. *Psychological bulletin,* 112 (1) pp.155-159.

Cole, T.J. 2002. A chart to link child centiles of body mass index, weight and height. *European Journal of Clinical Nutrition,* (12) pp.1194-1199.

Cole, T.J., Freeman, V.J. and Preece, M.A. 1995. Body mass index reference curves for the UK. *Archives of Disease in Childhood,* 73 pp.25-29.

Colley, R., Gorber, S.C. and Tremblay, M.S. 2010. Quality control and data reduction procedures for accelerometry-derived measures of physical activity. *Health Reports,* 21 (1) pp.63-69.

Collings, P.J., Brage, S., Ridgway, C.L., Harvey, N.C., Godfrey, K.M., Inskip, H.M., Cooper, C., Wareham, N.J. and Ekelund, U. 2013. Physical activity intensity, sedentary time, and body composition in preschoolers. *The American Journal of Clinical Nutrition,* 97 (5) pp.1020-1028.

Corder, K., Brage, S., Ramachandran, A., Snehalatha, C., Wareham, N. and Ekelund, U. 2007. Comparison of two Actigraph models for assessing free-living physical activity in Indian adolescents. *Journal of Sports Sciences,* 25 (14) pp.1607-1611.

Corder, K., Ekelund, U., Steele, R.M., Wareham, N.J. and Brage, S. 2008. Assessment of physical activity in youth. *Journal of Applied Physiology,* 105 (3) pp.977-987.

Corder, K., van Sluijs, E.M., Wright, A., Whincup, P., Wareham, N.J. and Ekelund, U. 2009. Is it possible to assess free-living physical activity and energy expenditure in young people by self-report? *American Journal of Clinical Nutrition,* 89 (3) pp.862-870.

Cox, R., Skouteris, H., Rutherford, L., Fuller-Tyszkiewicz, M., Dell' Aquila, D. and Hardy, L.L. 2012. Television viewing, television content, food intake, physical activity and body mass index: a cross-sectional study of preschool children aged 2-6 years. *Health Promotion Journal of Australia,* 23 (1) pp.58-62.

Crouter, S.E., Schneider, P.L., Karabulut, M. and Bassett, D.R. 2003. Validity of 10 electronic pedometers for measuring steps, distance, and energy cost. *Medicine and Science in Sports and Exercise,* 35 (8) pp.1455-1460.

Danner, F., Noland, M., McFadden, M., DeWalt, K. and Kotchen, J.M. 1991. Description of the physical activity of young children using movement sensor and observation methods. *Pediatric Exercise Science,* 3 (1) pp.11-20.

Davies, G., Reilly, J.J., McGowan, A.J., Dall, P.M., Granat, M.,H. and Paton, J.Y. 2012a. Validity, practical utility, and reliability of the activPAL™ in preschool children. *Medicine and Science in Sports and Exercise,* 44 (4) pp.761-768.

Davies, G., Reilly, J.J. and Paton, J.Y. 2012b. Objective measurement of posture and posture transitions in the pre-school child. *Physiological Measurement*, 33 pp.1913-1921.

Davis, C.T. 1980. Metabolic cost of exercise and physical performance in children with some observations on external loading. *Pediatric Exercise Science,* 45 (2-3) pp.95-102.

De Bock, F., Menze, J., Becker, S., Litaker, D., Fischer, J. and Seidel, I. 2010. Combining accelerometry and HR for assessing preschoolers' physical activity. *Medicine and Science in Sports and Exercise,* 42 (12) pp.2237-2243.

De Decker, E., De Craemer, M., Santos-Lozano, A., Van Cauwenberghe, E., De Bourdeaudhij, I. and Cardon, G. 2013. Validity of the activPAL™ and the Actigraph monitors in preschoolers. *Medicine and Science in Sports and Exercise*, [Epub ahead of print] .

de Vries, S.I., Bakker, I., Hopman-Rock, M., Hirasing, R.A. and van, M.W. 2006. Clinimetric review of motion sensors in children and adolescents. *Journal of Clinical Epidemiology,* 59 (7) pp.670-680.

de Vries, S.I., Garre, F.G., Engbers, L.H., Hildebrandt, V.H. and Van, B.S. 2011. Evaluation of neural networks to identify types of activity using accelerometers. *Medicine and Science in Sports and Exercise,* 43 (1) pp.101-107.

de Vries, S.I., Van Hirtum, H.W., Bakker, I., Hopman-Rock, M., Hirasing, R.A. and van, M.W. 2009. Validity and reproducibility of motion sensors in youth: a systematic update. *Medicine and Science in Sports and Exercise,* 41 (4) pp.818-827.

Denham-Deal, T. 2005. Preschool children's accumulated and sustained physical activity. *Perceptual and Motor Skills,* 100 (2) pp.443-450.

Department of Health and Ageing., 1999. *National physical activity guidelines for Australians.* Canberra, Australia: Department of Health and Aging.

Department of Health and Children, Health Service Executive., 2009. *The National Guidelines on Physical Activity for Ireland.* Ireland: Department of Health and Children.

Department of Health, Physical Activity, Health Improvement and Protection., 2011. *Start Active, Stay Active. A report on physical activity for health from the four home countries' Chief Medical Officers.* [online].London: Department of Health [Accessed January 16 2012]. Available from: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_128209.

Dietz, W.H. 1997. Periods of risk in childhood for the development of adult obesity- what do we need to learn. *Journal of Nutrition,* 127 (9) pp.S1884-S1886.

Dolinsky, D.H., Brouwer, R.J.N., Evenson, K.R., Siega-Riz, A. and Østbye, T. 2011. Correlates of sedentary time and physical activity among preschool-aged children. *Preventing Chronic Disease,* 8 (6) 11, pp.A131-A131.

Dollman, J., Okely, A.D., Hardy, L., Timperio, A., Salmon, J. and Hills, A.P. 2009. A hitchhikers guide to assessing young people's physical activity: deciding what method to use. *Journal of Science and Medicine in Sport,* 12 (5) pp.518-525.

Dorsey, K., Herrin, J., Krumholz, H. and Irwin, M. 2009. The utility of shorter epochs in direct motion monitoring. *Research Quarterly for Exercise and Sport,* 80 (3) pp.460-468.

Dowda, M., Pate, R.R., Sallis, J. and Freedson, P.S. 1997. Accelerometer (CSA) count cut points for physical activity intensity ranges in youth. *Medicine and Science in Sports and Exercise,* 29 (5) pp.72.

Dowda, M., Pate, R.R., Trost, S.G., Almeida, M.J.C.A. and Sirard, J.R. 2004. Influences of preschool policies and practices on children's physical activity. *Journal of Community Health,* 29 (3) pp.183-196.

Duncan, J.S., Hopkins, W.G., Schofield, G. and Duncan, E.K. 2008. Effects of weather on pedometer-determined physical activity in children. *Medicine and Science in Sports and Exercise,* 40 (8) pp.1432-1438.

Duncan, S., White, K., Sa'ulilo, L. and Schofield, G. 2011. Convergent validity of a piezoelectic pedometer and an omnidirectional accelerometer for measuring children's physical activity. *Pediatric Exercise Science,* 23 pp.399-410.

DuRant, R.H., Baranowski, T., Davis, H., Thompson, W.O., Puhl, J., Greaves, K.A. and Rhodes, T. 1992. Reliability and variability of heart rate monitoring in 3-, 4-, or 5-yr-old children. *Medicine and Science in Sports and Exercise,* 24 (2) pp.265-271.

DuRant, R.H., Baranowski, T., Johnson, M. and Thompson, W.O. 1994. The relationship among television watching, physical activity, and body composition of young children. *Pediatrics,* 94 (4 Part 1) pp.449-455.

DuRant, R.H., Baranowski, T., Puhl, J., Rhodes, T., Davis, H., Greaves, K.A. and Thompson, W.O. 1993. Evaluation of the Children's Activity Rating Scale (CARS) in young children. *Medicine and Science in Sports and Exercise,* 25 (12) pp.1415-1421.

Dwyer, G.M., Baur, L.A. and Hardy, L.L. 2009. The challenge of understanding and assessing physical activity in preschool-age children: thinking beyond the framework of intensity, duration and frequency of activity. *Journal of Science and Medicine in Sport,* 12 (5) pp.534-536.

Dwyer, G.M., Hardy, L.L., Peat, J.K. and Baur, L.A. 2011. The validity and reliability of a home environment preschool-age physical activity questionnaire (Pre-PAQ). *The International Journal of Behavioral Nutrition and Physical Activity* [online] 8 pp.86. Available at: http://www.ijbnpa.org/content/pdf/1479-5868-8-86.pdf [Accessed April 08 2012].

Edwardson, C.L. and Gorely, T. 2010. Epoch length and its effect on physical activity intensity. *Medicine and Science in Sports and Exercise,* 42 (5) pp.928-934.

Eiberg, S., Hasselstrom, H., Grønfeldt, V., Froberg, K., Svensson, J. and Andersen, L.B. 2005. Maximum oxygen uptake and objectively measured physical activity in Danish children 6-7 years of age: the Copenhagen school child intervention study. *British Journal of Sports Medicine*, 39 (10) pp.725-730.

Eijkemans, M., Mommers, M., de Vries, S., van Buuren, S., Stafleu, A., Bakker, I. and Thijs, C. 2008. Asthmatic symptoms, physical activity, and overweight in young children: a cohort study. *Pediatrics,* 121 (3) pp.666-72.

Ekelund, U., Anderssen, S., Andersen, L.B., Riddoch, C.J., Sardinha, L.B., Luan, J., Froberg, K. and Brage, S. 2009. Prevalence and correlates of metabolic syndrome in a population-based sample of European youth. *The American Journal of Clinical Nurtition,* 89 (1) pp.90-96.

Ekelund, U., Brage, S., Froberg, K., Harro, M., Anderssen, S.A., Sardinha, L.B., Riddoch, C. and Andersen, L.B. 2006. TV viewing and physical activity are independently associated with metabolic risk in children: the European Youth Heart Study. *PLoS Medicine,* 3 (12) pp.e488.

Ekelund, U., Luan, J., Sherar, L.B., Esliger, D.W., Griew, P. and Cooper, A. 2012. Moderate to vigorous physical activity and sedentary time and cardiometabolic risk factors in children and adolescents. *Journal American Medical Association,* 307 (7) pp.704-712.

Ekelund, U., Sjöström, M., Yngve, A., Poortvliet, E., Nilsson, A., Froberg, K., Wedderkopp, N. and Westerterp, K. 2001. Physical activity assessed by activity monitor and doubly labeled water in children. *Medicine and Science in Sports and Exercise,* 33 (2) pp.275-281.

Ekelund, U., Tomkinson, G. and Armstrong, N. 2011. What proportion of youth are physically active? Measurement issues, levels and recent time trends. *British Journal of Sports Medicine,* 45 (11) pp.859-865.

Esliger, D.W., Copeland, J.L., Barnes, J.D. and Tremblay, M.S. 2005. Standardizing and optimizing the use of accelerometer data for free-living physical activity monitoring. *Journal of Physical Activity and Health,* 2 (3) pp.366-383.

Esliger, D.W. and Tremblay, M.S. 2006. Technical reliability assessment of three accelerometer models in a mechanical setup. *Medicine and Science in Sports and Exercise,* 38 (12) pp.2173-2181.

Eston, R.G., Rowlands, A.V. and Ingledew, D.K. 1998. Validity of heart rate, pedometry, and accelerometry for predicting the energy cost of children's activities. *Pediatric Exercise Science,* 84 (1) pp.362-371.

Evenson, K.R., Catellier, D.J., Gill, K., Ondrak, K.S. and McMurray, R.G. 2008. Calibration of two objective measures of physical activity for children. *Journal of Sports Sciences,* 26 (14) pp.1557-1565.

Evenson, K.R. and Terry Jr, J.W. 2009. Assessment of differing definitions of accelerometer nonwear time. *Research Quarterly for Exercise and Sport,* 80 (2) pp.355-362.

Fairweather, S.C., Reilly, J.J., Grant, S., Whittaker, A. and Paton, J.Y. 1999. Using the Computer Science and Applications (CSA) activity monitor in preschool children. *Pediatric Exercise Science,* 11 (4) pp.413-420.

Field, A. 2012. *Discovering Statistics Using SPSS.* London: Sage Publications.

Finan, K., Larson, D.E. and Goran, M.I. 1997. Cross-validation of prediction equations for resting energy expenditure in young, healthy children. *Journal of the American Dietetic Association,* 97 (2) pp.140-145.

Finn, K., Johannsen, N. and Specker, B. 2002. Factors associated with physical activity in preschool children. *Journal of Pediatrics*, 140 (1) pp.81-85.

Finn, K. and Ullmann, J. 2004. Does seasonal changes in temperature affect levels of vigorous physical activity in preschool-aged children? *Medicine and Science in Sports and Exercise,* 36 (Suppl 5) pp.S 298.

Finn, K.J. and Specker, B. 2000. Comparison of Actiwatch activity monitor and Children's Activity Rating Scale in children. *Medicine and Science in Sports and Exercise,* 32 (10) pp.1794-1797.

Firrincieli, V., Keller, A. and Eherensberger, R. 2005. Decreased physical activity among Head Start children with a history of wheezing: use of an accelerometer to measure activity. *Pediatric Pulmonology,* 40 pp.57-63.

Fisher, A., Reilly, J.J., Kelly, L.A., Montgomery, C., Williamson, A., Paton, J.Y. and Grant, S. 2005a. Fundamental movement skills and habitual physical activity in young children. *Medicine and Science in Sports and Exercise*, 37 (4) pp.684-688.

Fisher, A., Reilly, J.J., Montgomery, C., Kelly, L., Williamson, A., Jackson, D.M., Paton, J.Y. and Grant, S. 2005b. Seasonality in physical activity and sedentary behavior in young children. *Pediatric Exercise Science,* 17 (1) pp.31-40.

Fleiss, J.L. 1986. *The design and analysis of clinical experiments.* New York: Wiley.

Fontvieille, A.M., Harper, I.T., Ferraro, T.L., Spraul, M. and Ravussin, E. 1993. Daily energy expenditure by five-year-old children, measured by doubly labeled water. *Journal of Pediatrics*, 123 (2) pp.200-207.

Freedson, P., Bowles,H.R.,Troiano,R. and Haskell, W. 2012. Assessment of physical activity using wearable monitors: recommendations for monitor calibration and use in the field. *Medicine and Science in Sports and Exercise*, 44 (1) pp.S1-S4.

Freedson, P., Pober, D. and Janz, K.F. 2005. Calibration of accelerometer output for children. *Medicine and Science in Sports and Exercise*, 37 (Suppl 11) pp.S523-S530.

Freedson, P.S., Melanson, E. and Sirard, J. 1998. Calibration of the Computer Science and Applications, Inc. accelerometer. *Medicine and Science in Sports and Exercise,* 30 (5) pp.777-781.

Freedson, P.S. and Miller, K. 2000. Objective monitoring of physical activity using motion sensors and heart rate. *Research Quarterly for Exercise and Sport,* 71 (Suppl 2) pp.S21-S29.

Freedson, P.S., Sirard, J., Debold, E., Pate, R., Dowda, M., Trost, S. and Sallis, J. 1997. Calibration of the Computer Science and Applications, Inc. (CSA) accelerometer. *Medicine and Science in Sports and Exercise,* 29 (Suppl 5) pp.S45.

Fudge, B.W., Wilson, J., Easton, C., Irwin, L., Clark, J., Haddow, O., Kayser, B. and Pitsiladis, Y.P. 2007. Estimation of oxygen update during fast running using accelerometry and heart rate. *Medicine and Science in Sports and Exercise,* 39 (1) pp.192-198.

Gabel, L., Obeid, J., Nguyen, T., Proudfoot, N.A. and Timmons, B.W. 2011. Short-term muscle power and speed in preschoolers exhibit stronger tracking than physical activity. *Applied Physiology, Nutrition, and Metabolism,* 36 (6) pp.939-945.

Gabriel, K.P., McClain, J.J., Scmid, K.K., Storti, K.L., High, R.R., Underwood, D.A., Kuller, L.H. and Kriska, A.M. 2010. Issues in accelerometer methodology: the role of epoch length on estimates of physical activity and relationships with health outcomes in overweight, post-menopausal women . *International Journal of Behavioural Nutrition and Physical Activity* [online] 7 pp.53. Available at: http://www.ijbnpa.org/content/7/1/53 [Accessed June 3 2013].

Godfrey, A., Conway, R., Meagher, D. and ÓLaighin, G. 2008. Direct measurement of human movement by accelerometry. *Medical Engineering and Physics,* 30 (10) pp.1364-1386.

Godfrey, A., Culhane, K.M. and Lyons, G.M. 2007. Comparison of the performance of the activPAL^{TM} professional physical activity logger to discrete accelerometer-based activity monitor. *Medical Engineering and Physics,* 29 (8) pp.930-934.

Goldfield, G.S., Harvey, A., Grattan, K. and Adamo, K.B. 2012. Physical activity promotion in the preschool years: a critical period to intervene. *International Journal of Environmental Research and Public Health,* 9 (4) pp.1326-1342.

Goran, M.I. 1994. Application of the doubly labeled water technique for studying total energy expenditure in young children: a review. *Pediatric Exercise Science,* 6 (1) pp.11-30.

Goran, M.I., Hunter, G., Nagy, T.R. and Johnson, R. 1997. Physical activity related energy expenditure and fat mass in young children. *International Journal of Obesity,* 21 (3) pp.171-178.

Goran, M.I. and Treuth, M.S. 2001. Energy expenditure, physical activity, and obesity in children. *Pediatric Clinics of North America,* 48 (4) pp.931-953.

Grant, P.M., Ryan, C.G., Tigbe, W.W. and Granat, M.H. 2006. The validation of a novel activity monitor in the measurement of posture and motion during everyday activities. *British Journal of Sports Medicine,* 40 pp.992-997.

Guinhouya, B., Samouda, H., Zitouni, D., Vilhelm, C. and Hubert, H. 2011. Evidence of the influence of physical activity on metabolic syndrome and/or insulin resistance in the pediatric populations: A systematic review. *International Journal of Pediatric Obesity,* 6 (5-6) pp.361-388.

Guinhouya, C.B. and Hubert, H. 2008. Incoherence with studies using Actigraph MTI among children age 6-12 yr. *Medicine and Science in Sports and Exercise,* 40 (5) pp.979.

Guinhouya, C.B., Hubert, H., Soubrier, S., Vilhelm, C., Lemdani, M. and Durocher, A. 2006. Moderate-to-vigorous physical activity among children: discrepancies in accelerometry-based cut-off points. *Obesity,* 14 (5) pp.774-777.

Hands, B., Parker, H. and Larkin, D. 2006. Physical activity measurement methods for young children: a comparative study. *Measurement in Physical Education and Exercise Science,* 10 (3) pp.203-214.

Harrell, J.S., McMurray, R.G., Baggett, C.D., Pennell, M.L., Pearce, P.F. and Bangdiwala, S.I. 2005. Energy costs of physical activity in children and adolescents. *Medicine and Science in Sports and Exercise,* 37 (2) pp.329-336.

Harro, M. 1997. Validation of a questionnaire to assess physical activity of children aged 4-8 years. *Research Quarterly for Exercise and Sport,* 68 (4) pp.259-268.

Health Canada., 2002. *Canada's physical activity guide for children.* Ottawa, ON: Health Canada.

Healy, G.N., Mathews, C.E., Dunstan, D.W., Winkler, E.A.H. and Owen, N. 2011. Sedentary time and cardio-metabolic biomarkers in US adults: NHANES 2003-06. *European Heart Journal,* 32 (2) pp.590-597.

Heelan, K.A. and Eisenmann, J.C. 2006. Physical activity, media time, and body composition in young children. *Journal of Physical Activity and Health,* 3 (2) pp.200-209.

Hendelman, D., Miller, K., Baggett, C., Debold, E. and Freedson, P. 2000. Validity of accelerometry for the assessment of moderate intensity physical activity in the field. *Medicine and Science in Sports and Exercise,* 32 (Suppl 9) pp.S442-S449.

Hendrick, P., Boyd, T., Low, O., Takarangi, K., Paterson, M., Claydon, L. and Milosavljevic, S. 2010. Construct validity of RT3 accelerometer: a comparison of level-ground and treadmill walking at self-selected speeds. *Journal of Rehabilitation Research and Development,* 47 (2) pp.157-168.

Hinkley, T., Salmon, J., Okely, A.D., Hesketh, K. and Crawford, D. 2012a. Correlates of preschool children's physical activity. *American Journal of Preventive Medicine,* 43 (2) pp.159-167.

Hinkley, T., O'Connell, E., Okely, A., D., Crawford, D., Hesketh, K. and Salmon, J. 2012b. Assessing volume of accelerometry data for reliability in preschool children. *Medicine and Science in Sports and Exercise,* 44 (12) pp.2436-2441.

Hinkley, T., Salmon, J., Okely, A., D., Crawford, D. and Hesketh, K. 2012c. Preschoolers' physical activity, screen time, and compliance with recommendations. *Medicine and Science in Sports and Exercise,* 44 (3) pp.458-465.

Hislop, J.F., Bulley, C., Mercer, T.H. and Reilly, J.J. 2012a. Comparison of accelerometry cut points for physical activity and sedentary behavior in preschool children: a validation study. *Pediatric Exercise Science*, 24 (4) pp.563-576.

Hislop, J.F., Bulley, C., Mercer, T.H. and Reilly, J.J. 2012b. Comparison of epoch and uniaxial versus triaxial accelerometers in the measurement of physical activity in preschool children: a validation study. *Pediatric Exercise Science,* 24 (3) pp.450-460.

Hoos, M.B., Kuipers, H., Gerver, W.J. and Westerterp, K.R. 2004. Physical activity pattern of children assessed by triaxial accelerometry. *European Journal of Clinical Nutrition,* 58 (10) pp.1425-1428.

Hoos, M.B., Plasqui, G., Gerver, W.J. and Westerterp, K.R. 2003. Physical activity level measured by doubly labeled water and accelerometry in children. *European Journal of Applied Physiology,* 89 (6) pp.624-626.

Hopkins, W.G. 2000. Measures of reliability in sports medicine and science. *Sports Medicine,* 30 (1) pp.1-15.

Hume, C., Salmon, J., Veitch, J., O'Connell, E., Crawford, D. and Ball, K. 2012. Socio-demographic characteristics of children experiencing socioeconomic disadvantage who meet physical activity and screen-time recommendations: the READI study. *Preventive Medicine,* 54 (1) pp.61-64.

Hussey, J., Bennett, K., Dwyer, J.O., Langford, S., Bell, C. and Gormley, J. 2009. Validation of the RT3 in the measurement of physical activity in children. *Journal of Science & Medicine in Sport,* 12 (1) pp.130-133.

Hussey, J., Gormley, J. and Bell, C. 2001. Physical activity in Dublin children aged 7-9 years. *British Journal of Sports,* 35 (4) pp.268-272.

Institute of Medicine of the National Academies. 2011. *Early childhood obestiy prevention policies.* Washington: The National Academic Press.

ISD Scotland. 2010. *Historical childhood Body Mass Index.* [online] Edinburgh, Scotland: Information Services Division, NHS National Services Scotland. Available at: http://www.isdscotland.org/child_obesity/Historical%20Childhood%20Body%20Mass%20Index%20(BMI)%20Statistics.pdf [Accessed June 12 2010].

Ivarsson, A., Andersen, M.B., Johnson, U. and Lindwall, M. 2013. To adjust or not adjust: Non parametric effect sizes, confidence intervals, and real-world meaning. *Psychology of Sports and Exercise,* 14 pp.97-102.

Jackson, D.M., Djafarian, K., Stewart, J. and Speakman, J.R. 2009. Increased television viewing is associated with elevated body fatness but not with lower total energy expenditure in children. *American Journal of Clinical Nutrition,* 89 (4) pp.1031-1036.

Jackson, D.M., Reilly, J.J., Kelly, L.A., Montgomery, C., Grant, S. and Paton, J.Y. 2003. Objectively measured physical activity in a representative sample of 3- to 4-year-old children. *Obesity Research,* 11 (3) pp.420-425.

Jago, R., Baranowski, T., Baranowski, J.C., Thompson, D. and Greaves, K.A. 2005. BMI from 3-6 y of age is predicted by TV viewing and physical activity, not diet. *International Journal of Obesity,* 29 (6) pp.557-564.

Janssen, I. and LeBlanc, A.G. 2010. Systemic review of the health benefits of physical activity and fitness in school-aged children and youth. *International Journal of Behavioral Nutrition and Physical Activity* [online] (7), pp.40. Available at: http://www.biomedcentral.com/content/pdf/1479-5868-7-40.pdf&embedded=true [Accessed July 25 2013].

Janz, K.F. 2006. Physical activity in epidemiology: moving from questionnaire to objective measurement. *British Journal of Sports Medicine,* 40 (3) pp.191-192.

Janz, K.F., Broffitt, B. and Levy, S. 2005. Validation evidence for the Netherlands physical activity questionnaire for young children: the Iowa bone development study. *Research Quarterly for Exercise and Sport,* 76 (3) pp.363-369.

Janz, K.F., Burns, T.L., Levy, S.M., Torner, J.C., Willing, M.C., Beck, T.J., Gilmore, J.M. and Marshall, T.A. 2004. Everyday activity predicts bone geometry in children: the Iowa bone development study. *Medicine and Science in Sports and Exercise,* 36 (7) pp.1124-1131.

Janz, K.F., Burns, T.L., Torner, J.C., Levy, S.M., Paulos, R., Willing, M.C. and Warren, J.J. 2001. Physical activity and bone measures in young children: the Iowa bone development study. *Pediatrics,* 107 (6) pp.1387-1393.

Janz, K.F., Letuchy, E.M., Eichenberger Gilmore, J.M., Burns, T.L., Torner, J.C., Willing, M.C. and Levy, S.M. 2010. Early physical activity provides sustained bone health benefits later in childhood. *Medicine and Science in Sports and Exercise,* 42 (6) pp.1072-1078.

Janz, K.F., Levy, S.M., Burns, T.L., Torner, J.C., Willing, M.C. and Warren, J.J. 2002. Fatness, physical activity and television viewing in children during the adiposity rebound period: the Iowa bone development study. *Preventive Medicine,* 35 (6) pp.536-571.

Jimmy, G., Dossegger, A., Seiler, R. and Mader, U. 2012. Metabolic thresholds and validated accelerometer cutoff points for the Actigraph GT1M in young children based on measurements of locomotion and play activities. *Measurement in Physical Education and Exercise Science,* 16 pp.23-40.

John, D. and Freedson, P. 2012. Actigraph and Actical physical activity monitors: a peek under the hood. *Medicine and Science in Sports and Exercise,* 44 (1 Suppl 1) pp.S86-S89.

John, D., Tyo, B. and Bassett, D.R. 2010. Comparison of four Actigraph accelerometers during walking and running. *Medicine and Science in Sports and Exercise,* 42 (2) pp.368-374.

Joy, J.E., Penhoet, E.E. and Petitti, D.E. eds. 2005. *Saving women's lives: strategies for improving breast cancer detection and diagnosis.* Washington (DC): National Academic Press (US).

Kahan, D., Nicaise, V. and Reuben, K. 2013. Convergent validity of four accelerometer cutpoints with direct observation of preschool children's outdoor physical activity. *Research Quarterly for Exercise and Sport,* 84 (1) pp.59-67.

Kambas, A., Michalopoulou, M., Fatouros, I., G., Christoforidis, C., Manthou, E., Giannakidou, D., Venetsanou, F., Haberer, E., Chatzinikolaou, A., Gourgoulis, V. and Zimmer, R. 2012. The relationship between motor proficiency and pedometer-determined physical activity in young children. *Pediatric Exercise Science,* 24 (1) pp.34-44.

Kaminsky, L.A. and Ozemek, C. 2012. A comparison of the Actigraph GT1M and GT3X accelerometers under standardized and free-living conditions. *Physiological Measurement,* 33 (11) pp.1869-1876.

Katzmarzyk, P.T. and Ardern, C.I. 2004. Physical activity levels of Canadian children and youth: current issues and recommendations. *Canadian Journal of Diabetes,* 28 (1) pp.67-78.

Kelly, L.A., Reilly, J.J., Fairweather, S.C., Barrie, S., Grant, S. and Paton, J.Y. 2004. Comparison of two accelerometers for assessment of physical activity in preschool children. *Pediatric Exercise Science,* 16 (4) pp.324-333.

Kelly, L.A., Reilly, J.J., Fisher, A., Montgomery, C., Williamson, A., McColl, J.H., Paton, J.Y. and Grant, S. 2006. Effect of socioeconomic status on objectively measured physical activity. *Archives of Disease in Childhood,* 91 (1) pp.35-38.

Kelly, L.A., Reilly, J.J., Grant, S. and Paton, J.Y. 2004. Objective measurement of physical activity in pre-school children: comparison of two accelerometers against direct observation. *Medicine and Science in Sports and Exercise,* 36 (Suppl 5) pp.S329.

Kelly, L.A., Reilly, J.J., Grant, S. and Paton, J.Y. 2005. Low physical activity levels and high levels of sedentary behaviour are characteristic of rural Irish primary school children. *Irish Medical Journal,* 98 (5) pp.138-141.

Kelly, L.A., Reilly, J.J., Jackson, D.M., Montgomery, C., Grant, S. and Paton, J.Y. 2007. Tracking physical activity and sedentary behavior in young children. *Pediatric Exercise Science,* 19 (1) pp.51-60.

Kim, Y., Beets, M.W. and Welk, G.J. 2012. Everything you wanted to know about selecting the "right" Actigraph accelerometer cut-point for youth, but....A systematic review. *Journal of Science and Medicine in Sport,* 15 (4) pp.311-321.

Kim, Y., Beets, M.W., Pate, R.R. and Blair, S.N. 2013. The effect of reintegrating Actigraph accelerometer counts in preschool children: comparison using different epoch lengths. *Journal of Science and Medicine in Sport,* 16 (2) pp.129-134.

King, W., Li, J., Leishear, K., Mitchell, J.E. and Belle, S. 2008. Accelerometer data reduction influences estimates of physical activity. *Obesity,* 16 (1) pp.S274.

Klesges, R.C., Coates, T.J. and Moldenhauer-Klesges, L.M. 1984. The FATS: an observational system for assessing physical activity in children and associated parent behavior. *Behavioral Assessment,* 6 (4) pp.333-345.

Kozey, S.L., Lyden, K., Howe, C.A., Staudenmayer, J.W. and Freedson, P.S. 2010a. Accelerometer output and MET values of common physical activities. *Medicine and Science in Sports and Exercise*, 42 (9) pp.1776-1784.

Kozey, S.L., Staudenmayer, J.W., Troiano, R.P. and Freedson, P.S. 2010b. Comparison of the Actigraph 7164 and the Actigraph GT1M during self-paced locomotion. *Medicine and Science in Sports and Exercise*, 42 (5) pp.971-976.

Kozey-Keadle, S., Lyden, K., Staudenmayer, J. and Freedson, P. 2010. Errors in MET estimates of physcial activity using 3.5ml·Kg^{-1}·min^{-1} as the baseline oxygen consumption. *Journal of Physical Activity and Health*, 7 (4) pp.508-516.

Kriska, A.M., Knowler, W.C., LaPorte, R.E., Drash, A.L., Wing, R.R., Blair, S.N., Bennett, P.H. and Kuller, L.H. 1990. Development of questionnaire to examine the relationship of physical activity and diabetes in Pima Indians. *Diabetes Care*, 13 (4) pp.401-411.

Kristensen, P.L., Moller, N.C., Korsholm, L., Wedderkopp, N., Andersen, L.B. and Froberg, K. 2008. Tracking of objectively measured physical activity from childhood to adolescence: the European youth heart study. *Scandinavian Journal of Medicine and Science in Sports*, 18 (2) pp.171-178.

Kwon, S. and Janz, K.F. 2012. Tracking of accelerometry-measured physical activity during childhood: ICAD pooled analysis. *International Journal of Behavioural Nutrition and Physical Activity* [online] 68 (9), Available at: http://www.ijbnpa.org/content/pdf/1479-5868-9-68.pdf [Accessed March 12 2013].

Lambourne, K. and Donnelly, J.E. 2011. The role of physical activity in pediatric obesity. *Pediatric Clinics of North America*, 58 (6) pp.1481-1491.

Landis, J.R. and Koch, G.G. 1977. The measurement of observer agreement for categorical data. *Biometrics*, 33 (1) pp.159-174.

Lanigan, J., Barber, S. and Singhal, A. 2010. Prevention of obesity in preschool children. *Proceedings of the Nutrition Society*, 69 (2) pp.204-210.

Levine, J.A. 2005. Measurement of energy expenditure. *Public Health Nutrition*, 8 (7A) pp.1123-1132.

Levine, J.A., Lanningham-Foster, L.M., McCrady, S.K., Krizan, A.C., Olson, L.R., Kane, P.H., Jensen, M.D. and Clark, M.M. 2008. Interindividual variation in posture allocation: possible role in human obesity. *Science*, 307 (5709) pp.584-554.

Logan, N., Reilly, J.J., Grant, S. and Paton, J.Y. 2000. Resting heart rate definition and its effect on apparent levels of physical activity in young children. *Medicine and Science in Sports and Exercise*, 32 (1) pp.162-166.

Lopez-Alarcon, M., Merrifield, J., Fields, D.A., Hilario-Hailey, T., Franklin, F.A., Shewchuk, R.M., Oster, R.A. and Gower, B.A. 2004. Ability of the Actiwatch accelerometer to predict free-living energy expenditure in young children. *Obesity Research*, 12 (11) pp.1859-1865.

Loprinzi, P.D. and Cardinal, B.J. 2011. Measuring children's physical activity and sedentary behaviors. *Journal of Science and Fitness*, 9 (1) pp.15-23.

Louie, L. and Chan, L. 2003. The use of pedometry to evaluate the physical activity levels among preschool children in Hong Kong. *Early Child Development and Care*, 173 (1) pp.97-107.

Louie, L., Eston, R.G., Rowlands, A.V., Kwok, K.T., Ingledew, D.K. and Fu, F.H. 1999. Validity of heart rate, pedometry, and accelerometry for estimating the energy cost of activity in Hong Kong Chinese boys. *Pediatric Exercise Science*, 11 (3) pp.229-239.

Lubans, D.R., Hesketh, K., Cliff, D.P., Barnett, L.M., Salmon, J., Dollman, J., Morgan, P.J., Hills, A.P. and Hardy, L.L. 2011. A systematic review of the validity and reliability of sedentary behaviour measures used with children and adolescents. *Obesity Reviews*, 12 (10) pp.781-799.

Mackintosh, K.A., Fairclough, S.J., Stratton, G. and Ridgers, N.D. 2012. A calibration protocol for population-specific accelerometer cut-points in children. *PLoS ONE* [online] 7 (5), pp.e36919. Available at: http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0036919 [Accessed July 4 2013].

Mahar, M., Smith, S., Rowe, D., DuBose, K. and McCammon, M. 2008. Effect of sampling interval on objectively-measured physical activity in pre-school children. *Medicine and Science in Sports and Exercise*, 40 (Suppl 5) pp.S205.

Malina, R.M. 1996. Tracking of physical activity and physical fitness across the lifespan. *Research Quarterly for Exercise and Sport*, 67 (Suppl 3) pp.S48-S57.

Manufacturing Technology Inc., 2001. *Activity monitor operator's manual, software RIU 256k, version 2.26. The model AM7164 (CSA) monitor.* Fort Walton Beach, Florida: Manufacturing Technology, Inc.

Martin, A., McNeill, M., Penpraze, V., Dall, P., Granat, M., Paton, J.Y. and Reilly, J. 2011. Objective measurement of habitual sedentary behavior in pre-school children: comparison of activPAL with Actigraph monitors. *Pediatric Exercise Science*, 23 (4) pp.468-476.

Martinez-Gomez, D., Tucker, J., Heelan, K.A., Welk, G.J. and Eisenmann, J.C. 2009. Associations between sedentary behavior and blood pressure in young children. *Archives of Pediatrics and Adolescent Medicine*, 163 (8) pp.724-730.

Mâsse, L.C., Fuemmeler, B.F., Andersen, C.B., Matthews, C.E., Trost, S.G., Catellier, D.J. and Treuth, M. 2005. Accelerometer data reduction: a comparison of four reduction algorithms on selected outcome variables. *Medicine and Science in Sports and Exercise*, 37 (Suppl 11) pp.S544-S554.

Masse, L.C., Fulton, J.E., Watson, K.L., Heesch, K.C., Kohl, H.W., Blair, S.N. and Tortolero, S.R. 1999. Detecting bouts of physical activity in a field setting. *Research Quarterly for Exercise and Sport*, 70 (3) pp.212-219.

Matthews, C.E. 2005. Calibration of accelerometer output for adults. *Medicine and Science in Sports and Exercise,* 37 (Suppl 11) pp.S512-S522.

Matthews, C.E., Chen, K.Y., Freedson, P.S., Buchowski, M.S., Beech, B.M., Pate, R.R. and Troiano, R.P. 2008. Amount of time spent in sedentary behaviors in the United States, 2003-2004. *American Journal of Epidemiology,* 167 (7) pp.875-881.

Matthews, C.E., Hagströmer, M., Pober, D.M. and Bowles, H.R. 2012. Best practices for using physical activity monitors in population-based research. *Medicine and Science in Sports and Exercise,* 44 (1 Suppl 1) pp.S68-S76.

Mattocks, C., Leary, S., Ness, A., Deere, K., Saunders, J., Tilling, K., Kirkby, J., Blair, S.N. and Riddoch, C. 2007. Calibration of an accelerometer during free-living activities in children. *International Journal of Pediatric Obesity,* 2 (4) pp.218-226.

Mattocks, C., Ness, A., Leary, S., Tilling, K., Blair,S,N, Shield, J., Deere, K., Saunders, J., Kirkby, J., Smith, G.D., Wells, J., Wareham, N., Reilly, J. and Riddoch, C. 2008. Use of accelerometers in a large field-based study of children: protocols, design issues, and effects on precision. *Journal of Physical Activity and Health,* 5 (Suppl 1) pp.S98-S111.

McClain, J.J., Abraham, T.L., Brusseau, T.A.,Jr. and Tudor-Locke, C. 2008. Epoch length and accelerometer outputs in children: comparison to direct observation. *Medicine and Science in Sports and Exercise,* 40 (12) pp.2080-2087.

McGrath, L. and Hinckson, E. 2009. Classification of physical activity in children using accelerometers. *Journal of Science and Medicine in Sport,* 12 (Suppl) pp.S8-S8.

McIver, K.L., Pfeiffer, K.A., Dowda, M., Joao, M., Almeida, C.A. and Pate, R. 2005. Validity and inter-instrument reliability of the Actigraph accelerometer in 3-5 year old children. *Medicine and Science in Sports and Exercise,* 37 (Suppl 5) pp.S113.

McKee, D.P., Boreham, C.A.G., Murphy, M.H. and Nevill, A.M. 2005. Validation of the Digiwalker pedometer for measuring physical activity in young children. *Pediatric Exercise Science,* 17 (4) pp.345-352.

McKenzie, T.L. 2002. Use of direct observation to assess physical activity. In: Welk, G.J. ed. *Physical activity assessment for health-related research.* Champaign, IL: Human Kinetics, pp. 179-195.

McKenzie, T.L., Marshall, S.J., Sallis, J.F. and Conway, T.L. 2000. Leisure-time physical activity in school environments: an observation study using SOPLAY. *Preventive Medicine,* 30 (1) pp.70-77.

McKenzie, T.L., Sallis, J.F., Nader, P.R., Patterson, T.L., Elder, J.P., Berry, C.C., Rupp, J.W., Atkins, C.J., Buono, M.J. and Nelson, J.A. 1991. BEACHES: an observational system for assessing children's eating and physical activity behaviors and associated events. *Journal of Applied Behavior Analysis,* 24 (1) pp.141-151.

McNamara, E., Hudson, Z. and Taylor, S.J.C. 2010. Measuring activity levels of young people: the validity of pedometers. *British Medical Bulletin,* 95 pp.121-137.

Melanson, E.L. and Freedson, P.S. 1996. Physical activity assessment: a review of methods. *Critical Reviews in Food Science and Nutrition*, 36 (5) pp.385-396.

Metallinos-Katsaras, E.S., Freedson, P., Fulton, J.E. and Sherry, B. 2007. The association between an objective measure of physical activity and weight status in preschoolers. *Obesity*, 15 (3) pp.686-694.

Metcalf, B.S., Hosking, J., Jeffery, A.N., Voss, L.D., Henley, W. and Wilkin, T.J. 2011. Fatness leads to inactivity, but inactivity does not lead to fatness: a longitudinal study in children (EarlyBird 45). *Archives of Disease in Childhood*, 96 (10) pp.942-947.

Metcalf, B.S., Voss, L.D., Hosking, J., Jeffery, A.N. and Wilkin, T.J. 2008. Physical activity at the government-recommended level and obesity-related health outcomes: a longitudinal study (Early Bird 37). *Archives of Disease in Childhood*, 93 (9) pp.772-777.

Metcalf, B.S., Voss, L.D. and Wilkin, T.J. 2002. Accelerometers identify inactive and potentially obese children (EarlyBird 3). *Archives of Disease in Childhood*, 87 (2) pp.166-167.

Mitchell, J.A., Pate, R.R., Dowda, M., Mattocks, C., Riddoch, C., Ness, A.R. and Blair, S.N. 2012. A prospective study of sedentary behavior in a large cohort of youth. *Medicine and Science in Sports and Exercise*, 44 (6) pp.1081-1087.

Moller, N.C., Kristensen, P.L., Wedderkopp, N., Andersen, L.B. and Froberg, K. 2009. Objectively measured habitual physical activity in 1997/1998 vs 2003/2004 in Danish children: The European Youth Heart Study. *Scandinavian Journal of Medicine and Science in Sports*, 19 (1) pp.19-29.

Montgomery, C., Reilly, J.J., Jackson, D.M., Kelly, L.A., Slater, C., Paton, J.Y. and Grant, S. 2004. Relation between physical activity and energy expenditure in a representative sample of young children. *American Journal of Clinical Nutrition*, 80 (3) pp.591-596.

Moore, L.L., Nguyen, U.S., Rothman, K.J., Cupples, L.A. and Ellison, R.C. 1995. Preschool physical activity level and change in body fatness in young children. The Framingham Children's Study. *American Journal of Epidemiology*, 142 (9) pp.982-988.

Mota, J., Valente, M., Aires, L., Silva, P., Santos, M.P. and Ribeiro, J.C. 2007. Accelerometer cut-points and youth physical activity prevalence. *European Physical Education Review*, 13 (3) pp.287-299.

MRC epidemiology unit. 2013. *Physical activity downloads: MAHUffe programme.* [online].http://www.mrc-epid.cam.ac.uk/research/resources/materials-transfer-disclaimer/physical-activity-downloads/ ed. Cambridge: University of Cambridge.

Muller, M.J. and Bosy-Westphal, A. 2003. Assessment of energy expenditure in children and adolescents. *Current Opinion in Clinical Nutrition and Metabolic Care*, 6 (5) pp.519-530.

Munro, B.H. 2001. *Statistical Methods for Health Care Research.* Philadelphia: Lippincott.

Nakagawa, S. and Cuthill, I.C. 2007. Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological Reviews*, 82 (4) pp.591-605.

National Association for Sport and Physical Education., 2002. *Active start: a statement of physical activity guidelines for children from birth to five years.* Reston, VA: National Association for Sport and Physical Education Publications.

National Association for Sport and Physical Education., 2009. *Active Start: a statement of physical activity guidelines for children from birth to age 5.* 2nd ed. Reston, VA: American Alliance for Health, Physical Education, Recreation and Dance.

National Institute for Health and Clinical Excellence., 2009. *PH 17 Promoting physical activity for children and young people: quick reference guide.* [online].London: National Institute for Health and Clinical Excellence.

Nilsson, A., Brage, S., Riddoch, C., Anderssen, S., Sardinha, L., Wedderkopp, N. and Andersen, L.E.,U. 2008. Comparison of equations for predicting energy expenditure from accelerometer counts in children. *Scandinavian Journal of Medicine and Science in Sports,* 18 (5) pp.643-650.

Nilsson, A., Ekelund, U., Yngve, A. and Sjöström, M. 2002. Assessing physical activity among children with accelerometers using different time sampling intervals and placements. *Pediatric Exercise Science,* 14 (1) pp.87-96.

Nishikido, N., Kashiwazaki, H. and Suzuki, T. 1982. Preschool children's daily activities: direct observation, pedometry or questionnaire. *Journal of Human Ergology,* 11 (2) pp.214-218.

Noland, M., Danner, F., DeWalt, K., McFadden, M. and Kotchen, J.M. 1990. The measurement of physical activity in young children. *Research Quarterly for Exercise and Sport,* 61 (2) pp.146-153.

Nyberg, G., Ekelund, U. and Marcus, C. 2009. Physical activity in children measured by accelerometry: stability over time. *Scandinavian Journal of Medicine and Science in Sports,* 19 (1) pp.30-35.

Obeid, J., Nguyen, T., Gabel, L. and Timmons, B.W. 2011. Physical activity in Ontario preschoolers: prevalence and measurement issues. *Applied Physiology, Nutrition, and Metabolism.,* 36 (2) pp.291-297.

O'Dwyer, M.V., Foweather, L., Stratton, G. and Ridgers, N.D. 2011. Physical activity in non-overweight and overweight UK preschool children: preliminary findings and methods of the Active Play Project. *Science and Sports,* 26 (6) pp.345-349.

Office of Population Censuses and Surveys. Social Survey Division., 1995. *National diet, nutrition and dental survey children aged 1 1/2 to 4 1/2, 1992-1993.* [online].Department of Health.

O'Hara, N.M., Baranowski, T., Simons-Morton, B.G., Wilson, B.S. and Parcel, G. 1989. Validity of the observation of children's physical activity. *Research Quarterly for Exercise and Sport,* 60 (1) pp.42-47.

Ojiambo, R., Cuthill, R., Budd, H., Konstabel, K., Casajus, J.A., González-Agüero, A., Anjila, E., Reilly, J.J., Easton, C., Pitsiladis, Y.P. and IDEFICS Consortium. 2011. Impact of methodological decisions on accelerometer outcome variables in young children. *International Journal of Obesity*, 35 (Suppl 1) pp.S98-S103.

Ojiambo, R., Konstabel, K., Veidebaum, T., Reilly, J., Verbestel, V., Huybrechts, I., Sioen, I., Casajús, J.A., Moreno, L.A., Vicente-Rodriguez, G., Bammann, K., Tubic, B.M., Marild, S., Westerterp, K., Pitsiladis, Y.P. and IDEFICS Consortium. 2012. Validity of hip-mounted uniaxial accelerometry with heart-rate monitoring vs triaxial accelerometry in the assessment of free-living energy expenditure in young children: the IDEFICS validation study. *Journal of Applied Physiology,* 113 (10) pp.1530-1536.

Okely, A.D., Trost, S.G., Steele, J.R., Cliff, D.P. and Mickle, K. 2009. Adherence to physical activity and electronic media guidelines in Australian pre-school children. *Journal of Paediatrics and Child Health,* 45 pp.5-8.

Oliver, M., Badland, H., Schofield, G.M. and Shepard, J. 2012. Identification of accelerometer nonwear time and sedentary behavior. *Research Quarterly for Exercise and Sport,* 82 (4) pp.779-783.

Oliver, M., Schluter, J. and Schofield, G. 2011. A new approach for the analysis of accelerometery data measured on preschool children. *Journal of Physical Activity Health,* 8 (2) pp.296-304.

Oliver, M., Schofield, G. and Schluter, J. 2009. Accelerometry to assess preschooler's free play: issues with count thresholds and epoch durations. *Measurement in Physical Education and Exercise Science,* 13 (4) pp.181-190.

Oliver, M., Schofield, G.M., Kolt, G. and Schulter, P.J. 2007a. Pedometer accuracy in physical activity assessment of preschool children. *Journal of Science and Medicine in Sports,* 10 (5) pp.303-310.

Oliver, M., Schofield, G.M. and Kolt, G.S. 2007b. Physical activity in preschoolers: understanding prevalence and measurement issues. *Sports Medicine,* 37 (12) pp.1045-1070.

Oortwijn, A.W., Plasqui, G., Reilly, J.J. and Okely, A.D. 2009. Feasibility of an activity protocol for young children in a whole room indirect calorimeter: a proof-of-concept study. *Journal of Physical Activity and Health,* 6 (5) pp.633-637.

Ott, A.E., Pate, R.R., Trost, S.G., Ward, D.S. and Saunders, R. 2000. The use of uniaxial and triaxial accelerometers to measure children's "free-play" physical activity. *Pediatric Exercise Science,* 12 (4) pp.360-370.

Parsons, T.J., Power, C., Logan, S. and Summerbell, C.D. 1999. Childhood predictors of adult obesity: a systematic review. *International Journal of Obesity,* 23 Suppl 8 pp.S1-107.

Pate, R.P., Mitchell, J.A., Byun, W. and Dowda, M. 2011. Sedentary behaviour in youth. *British Journal of Sports Medicine,* 45 (11) pp.906-913.

Pate, R.R., Almeida, M.J., McIver, K.L., Pfeiffer, K.A. and Dowda, M. 2006. Validation and calibration of an accelerometer in preschool children. *Obesity,* 14 (11) pp.2000-2006.

Pate, R.R., O'Neill, J.R. and Lobelo, F. 2008. The evolving definition of "sedentary". *Exercise and Sport Sciences Reviews,* 36 (4) pp.173-178.

Pate, R.R., O'Neill, J.R. and Mitchell, J. 2010. Measurement of physical activity in preschool children. *Medicine and Science in Sports and Exercise,* 42 (3) pp.508-512.

Pate, R.R., Trost, S.G., Dowda, M., Ott, A.E., Ward, D.S., Saunders, R. and Felton, G. 1999. Tracking of physical activity, physical inactivity, and health-related physical fitness in rural youth. *Pediatric Exercise Science,* 11 (4) pp.364-376.

Payne, G. and Isaacs, D. 1995. *Human Motor Development: A Lifespan Approach.* London: Mayfield Publishing Company.

Pellegrini, A.D. and Smith, P.K. 1998. Physical activity play: the nature and function of a neglected aspect of playing. *Child Development,* 69 (3) pp.577-598.

Pellegrini, A.D. 2010. Games and play mean different things in an educational context. *Nature,* 467 (7311) pp.27-27.

Penpraze, V., Reilly, J.J., Maclean, C.M., Montgomery, C., Kelly, L.A., Paton, J.Y., Aitchison, T. and Grant, S. 2006. Monitoring of physical activity in young children: how much is enough? *Pediatric Exercise Science,* 18 (4) pp.483-491.

Pettee Gabriel, K.K., Morrow, J.R.,Jr. and Woolsey, A.L. 2012. Framework for physical activity as a complex and multidimensional behavior. *Journal of Physical Activity and Health,* 9 (Suppl 1) pp.S11-S18.

Pfeiffer, K.A., Dowda, M., McIver, K.L. and Pate, R.R. 2009. Factors related to objectively measured physical activity in preschool children. *Pediatric Exercise Science,* 21 (2) pp.196-208.

Pfeiffer, K.A., McIver, K.L., Dowda, M., Almeida, M.J. and Pate, R.R. 2006. Validation and calibration of the Actical accelerometer in preschool children. *Medicine and Science in Sports and Exercise,* 38 (1) pp.152-157.

Pfeiffer, K.A., McIver, K.L., Dowda, M., Joao, M., Almeida, C.A. and Pate, R. 2005. Calibration and cross-validation of the Actical accelerometer in preschool aged children: 589 Board 180 2:00PM - 3:00PM. *Medicine and Science in Sports and Exercise,* 37 (Suppl) pp.S114.

Pober, D.M., Staudenmayer, J., Raphael, C. and Freedson, P.S. 2006. Development of novel techniques to classify physical activity mode using accelerometers. *Medicine and Science in Sports and Exercise,* 38 (9) pp.1626-1634.

Powell, S.M. and Rowlands, A.V. 2004. Intermonitor variability of the RT3 accelerometer during typical physical activities. *Medicine and Science in Sports and Exercise,* 36 (2) pp.324-330.

Puhl, J., Greaves, K., Hoyt, M. and Baranowski, T. 1990. Children's Activity Rating Scale (CARS): description and calibration. *Research Quarterly for Exercise and Sport,* 61 (1) pp.26-36.

Pulsford, R.M., Cortina-Borja, M., Rich, C., Kinnafick, F., Dezateux, C. and Griffiths, L.J. 2011. Actigraph accelerometer-defined boundaries for sedentary behavior and physical intensities in 7 year old children. *PLos ONE* [online] 6 (8), pp.e21822. Available at: http://www.plosone.org/article/fetchObject.action?uri=info%3Adoi%2F10.1371%2Fjournal.pone.0021822&representation=PDF [Accessed March 3 2012].

Puyau, M.R., Adolph, A.L., Vohra, F.A. and Butte, N.F. 2002. Validation and calibration of physical activity monitors in children. *Obesity Research*, 10 (3) pp.150-157.

Puyau, M.R., Adolph, A.L., Vohra, F.A., Zakeri, I. and Butte, N.F. 2004. Prediction of activity energy expenditure using accelerometers in children. *Medicine and Science in Sports and Exercise*, 36 (9) pp.1625-1631.

Quigg, R., Gray, A., Reeder, A.I., Holt, A. and Waters, D.L. 2010. Using accelerometers and GPS units to identify the proportion of daily physical activity located in parks with playgrounds in New Zealand children. *Preventive Medicine,* 50 (5-6) pp.235-240.

Raudsepp, L. and Pall, P. 1998. Reproducibility and stability of physical activity in children. *Pediatric Exercise Science,* 10 (4) pp.320-326.

Reilly, J.J. 2010. Low levels of objectively measured physical activity in preschoolers in child care. *Medicine and Science in Sports and Exercise,* 42 (3) pp.502-507.

Reilly, J.J., Coyle, J., Kelly, L., Burke, G., Grant, S. and Paton, J.Y. 2003. An objective method for measurement of sedentary behavior in 3- to 4-year olds. *Obesity Research,* 11 (10) pp.1155-1158.

Reilly, J.J., Jackson, D.M., Montgomery, C., Kelly, L.A., Slater, C., Grant, S. and Paton, J.Y. 2004. Total energy expenditure and physical activity in young Scottish children: mixed longitudinal study. *Lancet,* 363 (9404) pp.211-212.

Reilly, J.J., Kelly, L.A., Montgomery, C., Jackson, D.M., Slater, C., Grant, S. and Paton, J.Y. 2006a. Validation of Actigraph accelerometer estimates of total energy expenditure in young children. *International Journal of Pediatric Obesity,* 1 (3) pp.161-167.

Reilly, J.J., Kelly, L.A., Montgomery, C., Williamson, A., Fisher, A., McColl, J.H., Lo Conte, R., Paton, J. and Grant, S. 2006b. Physical activity to prevent obesity in young children: cluster randomised controlled trial. *British Medical Journal,* 333 pp.1041.

Reilly, J.J., Penpraze, V., Hislop, J., Davies, G., Grant, S. and Paton, J.Y. 2008. Objective measurement of physical activity and sedentary behaviour: review with new data. *Archives of Disease in Childhood*, 93 (7) pp.614-619.

Rice, M.H. and Howell, C.C. 2000. Measurement of physical activity, exercise, and physical fitness in children: issues and concerns. *Journal of Pediatric Nursing,* 15 (3) pp.148-156.

Riddoch, C.J., Bo Andersen, L., Wedderkopp, N., Harro, M., Klasson-Heggeb¢, L., Sardinha, L.B., Cooper, A.R. and Ekelund, U. 2004. Physical activity levels and patterns of 9- and 15-yr-old European children. *Medicine and Science in Sports and Exercise,* 36 (1) pp.86-92.

Riddoch, C.J., Mattocks, C., Deere, K., Saunders, J., Kirkby, J., Tiling, K., Leary, S.D., Blair, S.N. and Ness, A.R. 2007. Objective measurement of levels and patterns of physical activity. *Archives of Disease in Childhood,* 92 (11) pp.963-967.

Ridgers, N.D., Stratton, G. and McKenzie, T.L. 2010. Reliability and validity of the System for Observing Children's Activity and Relationships during Play (SOCARP). *Journal of Physical Activity and Health,* 7 (1) pp.17-25.

Ridley, K., Olds, T., Hands, B., Larkin, D. and Parker, H. 2009. Intra-individual variation in children's physical activity patterns: implications for measurement. *Journal of Science and Medicine in Sport,* 12 (5) pp.568-572.

Ridley, K. and Olds, T.S. 2008. Assigning energy costs to activities in children: a review and synthesis. *Medicine and Science in Sports and Exercise,* 40 (8) pp.1439-1446.

Ried-Larsen, M., Brønd, J.C., Brage, S., Hansen, B.H., Grydeland, M., Andersen, L.B. and Møller, N.C. 2012. Mechanical and free living comparison of four generations of the Actigraph activity monitor. *International Journal of Behavioral Nutrition and Physical Activity* [online] 9 pp.113. Available at: http://www.ijbnpa.org/content/pdf/1479-5868-9-113.pdf [Accessed November 10 2012].

Robusto, K.M. and Trost, S.G. 2012. Comparison of three generations of Actigraph™ activity monitors in children and adolescents. *Journal of Sports Sciences,* 30 (13) pp.1429-1435.

Roemmich, J.N., Epstein, L.H., Raja, S., Yin, L., Robinson, J. and Winiewicz, D. 2006. Association of access to parks and recreational facilities with the physical activity of young children. *Preventive Medicine,* 43 (6) pp.437-441.

Rokholm, B., Baker, J.L. and Sørensen, T.I.A. 2010. The levelling off of the obesity epidemic since the year 1999- a review of evidence and perspectives. *Obesity Reviews,* 11 (12) pp.835-846.

Rothney, M.P., Apker, G.A., Song, Y. and Chen, K.Y. 2008. Comparing the performance of three generations of Actigraph accelerometers. *Journal of Applied Physiology,* 105 pp.1091-1097.

Rothney, M.P., Neumann, M., Beziat, A. and Chen, K.Y. 2007. An artificial neural network model of energy expenditure using nonintegrated acceleration signals. *Journal of Applied Physiology,* 103 (4) pp.1419-1427.

Routen, A.C., Upton, D., Edwards, M.G. and Peters, D.M. 2012. Discrepancies in accelerometer measured physical activity in children due to cut-point non-equivalence and placement site. *Journal of Sports Sciences*, 30 (12) pp.1303-1310.

Rowe, P.J., Schuldheisz, J.M. and van der Mars, H. 1997. Validation of SOFIT for measuring physical activity of first- to eighth-grade students. *Pediatric Exercise Science,* 9 (2) pp.136-149.

Rowland, T.W. and Green, G.M. 1988. Physiological responses to treadmill exercise in females: adult-child differences. *Medicine and Science in Sports and Exercise,* 20 (5) pp.474-478.

Rowlands, A.V. 2007. Accelerometer assessment of physical activity in children: an update. *Pediatric Exercise Science,* 19 (3) pp.252-266.

Rowlands, A.V. and Eston, R.G. 2005. Comparison of accelerometer and pedometer measures of physical activity in boys and girls, ages 8-10 years. *Research Quarterly for Exercise and Sport,* 76 (3) pp.251-257.

Rowlands, A.V. and Eston, R.G. 2007. The measurement and interpretation of children's physical activity. *Journal of Sports Science and Medicine,* 6 pp.270-276.

Rowlands, A.V., Eston, R.G. and Ingledew, D.K. 1997. Measurement of physical activity in children with particular reference to the use of heart rate and pedometry. *Sports Medicine,* 24 (4) pp.258-272.

Rowlands, A.V., Eston, R.G. and Ingledew, D.K. 1998. Comparison of the validity of heart rate monitoring, simple pedometry, uniaxial accelerometry and triaxial accelerometry for the assessment of typical children's activities. *Journal of Sports Sciences*, 16 (1) pp.29-30.

Rowlands, A.V., Powell, S.M., Humphries, R. and Eston, R.G. 2006. The effect of accelerometer epoch on physical activity output measures. *Journal of Exercise Science and Fitness,* 4 (1) pp.52-58.

Rowlands, A.V., Thomas, P.W., Eston, R.G. and Topping, R. 2004. Validation of the RT3 triaxial accelerometer for the assessment of physical activity. *Medicine and Science in Sports and Exercise,* 36 (3) pp.518-524.

Saakslahti, A., Numminen, P., Helenius, H., Tammi, A. and Viikari, I. 2004. Physical activity as a preventative measure for coronary heart disease risk factors in early childhood. *Scandinavian Journal of Medicine and Science in Sports,* 14 (3) pp.143-149.

Saint-Maurice, P.F., Welk, G., Ihmels, M.A. and Kapfl, J.R. 2011. Validation of the SOPLAY direct observation tool with an accelerometry-based physical activity monitor. *Journal of Physical Activity and Health,* 8 (8) pp.1108-1116.

Sallis, J.F. 1991. Self-report measures of children's physical activity. *Journal of School Health,* 61 (5) pp.215-219.

Sallis, J.F. and Saelens, B.E. 2000. Assessment of physical activity by self-report: status, limitations, and future directions. *Research Quarterly for Exercise and Sport,* 71 (Suppl 2) pp.S1-14.

Sallo, M. and Silla, R. 1997. Physical activity with moderate to vigorous intensity in preschool and first-grade schoolchildren. *Pediatric Exercise Science,* 9 (1) pp.44-54.

Sarzynski, M.A., Eisenmann, J.C., Welk, G.J., Tucker, J., Glenn, K., Rothschild, M. and Heelan, K. 2010. ACE I/D genotype, habitual physical activity, and blood pressure in children. *Pediatric Exercise Science*, 22 (2) pp.301-313.

Schofield, W.N. 1985. Predicting basal metabolic rate, new standards and review of previous work. *Clinical Nutrition,* 39 (Suppl 1) pp.5-41.

Scottish Executive., 2003. *Let's make Scotland more active: a strategy for physical activity.* [online].Edinburgh: HMSO.

Scottish Government. 2010. *Scottish Index of Multiple Deprivation (2009).* [online] Edinburgh: Scottish Government. Available at: http://www.scotland.gov.uk/Topics/Statistics/SIMD/BackgroundMethodology [Accessed December 6th 2012].

Sedentary Behaviour Research Network. 2012. Letter to the Editor: standardized use of the terms "sedentary" and "sedentary behaviours". *Applied Physiology, Nutrition, and Metabolism,* 37 (3) pp.540-542.

Sharma, S.V., Chuang, R. and Skala, K. 2011. Measuring physcial activity in preschoolers: reliability and validity of the System for Observing Fitness Instruction time for Preschoolers (SOFIT-P). *Measurement in Physical Education and Exercise Science,* 15 pp.257-273.

Sigmund, E., De Ste Croix, M., Miklánková, L. and Frömel, K. 2007. Physical activity patterns of kindergarten children in comparison to teenagers and young adults. *European Journal of Public Health,* 17 (6) pp.646-651.

Sigmund, E., Sigmundová, D. and El Ansari, W. 2009. Changes in physical activity in pre-schoolers and first-grade children: longitudinal study in the Czech Republic. *Child: Care, Health and Development,* 35 (3) pp.376-382.

Sim, J. and Wright, C.C. 2005. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Physical Therapy,* 85 (3) pp.257-268.

Sirard, J.R. and Pate, R.R. 2001. Physical activity assessment in children and adolescents. *Sports Medicine,* 31 (6) pp.439-454.

Sirard, J.R., Trost, S.G., Pfeiffer, K.A., Dowda, M. and Pate, R.R. 2005. Calibration and evaluation of an objective measure of physical activity in preschool children. *Journal of Physical Activity and Health,* 2 (3) pp.345-357.

Skouteris, H., Dell'Aquila, D., Baur, L.A., Dwyer, G.M., McCabe, M.P., Ricciardelli, L.A. and Fuller-Tyszkiewicz, M. 2012. Physical activity guidelines for preschoolers: a call for research to inform public health policy. *The Medical Journal of Australia,* 196 (3) pp.174-177.

Sleap, M. and Tolfrey, K. 2001. Do 9- to 12 yr-old children meet existing physical activity recommendations for health? *Medicine and Science in Sports and Exercise,* 33 (4) pp.591-596.

Specker, B. and Binkley, T. 2003. Randomized trial of physical activity and calcium supplementation on bone mineral content in 3- to 5- year old children. *Journal of Bone and Mineral Research,* 18 (5) pp.885-892.

Spittaels, H., Van Cauwenberghe, E., Verbestel, V., De Meester, F., Van Dyck, D., Verloigne, M., Haerens, L., Deforche, B., Cardon, G. and De Bourdeaudhuij, I. 2012. Objectively measured sedentary time and physical activity across the lifespan: a cross-sectional study in four age groups. *International Journal of Behavioural Nutrition and Physical Activity* [online] 9 pp.149. Available at: http://www.ijbnpa.org/content/pdf/1479-5868-9-149.pdf [Accessed June 23 2013].

Stanley, J.C. and Angoff, W.H. 1971. Reliability. In: Thorndike, R.L. ed. *Educational Measurement.* Washington DC: American Council on Education, pp. 395.

Staudenmayer, J., Pober, D., Crouter, S.E., Bassett, D.R. and Freedson, P. 2009. An artificial neural network to estimate physical activity energy expenditure and identify physical activity type from an accelerometer. *Journal of Applied Physiology,* 107 (4) pp.1300-1307.

Steele, R.M., van Sluijs, E.M., Cassidy, A., Griffin, S.J. and Ekelund, U. 2009. Targeting sedentary time or moderate- and vigorous-intensity activity: independent relations with adiposity in a population-based sample of 10-y-old British children. *American Journal of Clinical Nutrition,* 90 (5) pp.1185-1192.

Strath, S.J., Pfeiffer, K.A. and Whitt-Glover, M.C. 2012. Accelerometer use with children, older adults, and adults with functional limitations. *Medicine and Science in Sports and Exercise,* 44 (Suppl 1) pp.S77-S85.

Strong, W.B., Malina, R.M., Blimkie, C.J., Daniels, S.R., Dishman, R.K., Gutin, B., Hergenroeder, A.C., Must, A., Nixon, P.A., Pivarnik, J.M., Rowland, T., Trost, S. and Trudeau, F. 2005. Evidence based physical activity for school-age youth. *Journal of Pediatrics,* 146 (6) pp.732-737.

Sun, D.X., Schmidt, G. and Teo-Koh, S.M. 2008. Validation of the RT3 accelerometer for measuring physical activity of children in simulated free-living conditions. *Pediatric Exercise Science,* 20 (2) pp.181-197.

Swartz, A.M., Strath, S.J., Bassett, D.R., O'Brien, W.L., King, G.A. and Ainsworth, B.E. 2000. Estimation of energy expenditure using CSA accelerometers at hip and wrist sites. *Medicine and Science in Sports and Exercise,* 32 (Suppl 9) pp.S450-S456.

Tanaka, C. and Tanaka, S. 2009. Daily physical activity in Japanese preschool children evaluated by triaxial accelerometry: the relationship between period of engagement in moderate-to-vigorous physical activity and daily step counts. *Journal of Physiological Anthropology,* 28 (6) pp.283-288.

Tanaka, C., Tanaka, S., Kawahara, J. and Midorikawa, T. 2007. Triaxial accelerometry for assessment of physical activity in young children. *Obesity,* 15 (5) pp.1233-1241.

Tanaka, C., Hikihara, Y., Ohkawara, K. and Tanaka, S. 2012. Locomotive and non-locomotive activity as determined by triaxial accelerometry and physical fitness in Japanese preschool children. *Pediatric Exercise Science,* 24 (3) pp.420-434.

Taylor, R.W., Murdoch, L., Carter, P., Gerrard, D.F., Williams, S.M. and Taylor, B.J. 2009. Longitudinal study of physical activity and inactivity in preschoolers: the FLAME study. *Medicine and Science in Sports and Exercise,* 41 (1) pp.96-102.

Telford, A., Salmon, J., Jolley, D. and Crawford, D. 2004. Reliability and validity of physical activity questionnaires for children: the Children's Leisure Activities Study Survey (CLASS). *Pediatric Exercise Science,* 16 (1) pp.64-78.

Telford, A., Salmon, J., Timperio, A. and Crawford, D. 2005. Examining physical activity among 5- to 6- and 10- to 12-year-old children: the Children's Leisure Activities Study. *Pediatric Exercise Science,* 17 (3) pp.266-280.

Terwee, C.B., Mokkink, L.B., van Popple,M.N.,Chinapaw,M.J., van Mechelen, W. and de Vet, H.C. 2010. Qualitative attributes and measurement properties of physical activity questionnaires: a checklist. *Sports Medicine*, 40 (7) pp.525-537.

Thompson, B. 2002. What future quantitative social science research could look like: confidence intervals for effect size. *Educational Researcher*, 31 (3) pp.25-32.

Timmons, B.W., Leblanc, A.G., Carson, V., Connor Gorber, S., Dillman, C., Janssen, I., Kho, M.E., Spence, J.C., Stearns, J.A. and Tremblay, M.S. 2012. Systematic review of physical activity and health in the early years (aged 0-4 years). *Applied Physiology, Nutrition, and Metabolism*, 37 (4) pp.773-792.

Timmons, B.W., Naylor, P.J. and Pfeiffer, K.A. 2007. Physical activity for preschool children--how much and how? *Canadian Journal Of Public Health,* 98 (Suppl 2) pp.S122-S134.

Toftager, M., Duncan, S., Oliver, M., Kristensen, L., Christiansen, L. and Troelsen, J. 2012. Physical activity in adolescents-accelerometer data reduction criteria: number of valid days, minimum daily wear time, and non-wear time. *Journal of Science and Medicine in Sport*, 15 (Suppl 1) pp.S65-S66.

Toschke, J.A., von Kries, R., Rosenfeld, E. and Toschke, A.M. 2007. Reliability of physical activity measures from accelerometry among preschoolers in free-living conditions. *Clinical Nutrition*, 26 (4) pp.416-420.

Tremblay, M.S., Colley, R.C., Saunders, T.J., Healy, G.N. and Owens, N. 2010. Physiological and health implications of a sedentary lifestyle. *Applied Physiology, Nutrition, and Metabolism*, 35 (6) pp.725-740.

Tremblay, M.S., LeBlanc, A.G., Kho, M.E., Saunders, T.J., Larouche, R., Colley, R.C., Goldfied, G. and Gorber, S.C. 2011. Systematic review of sedentary behaviour and health indicators in school-aged children and youth. *International Journal of Behavioral Nutrition and Physical Activity* [online] 8 (98), Available at: http://www.ijbnpa.org/content/pdf/1479-5868-8-98.pdf [Accessed May 21 2012].

Treuth, M.S., Schmitz, K., Catellier, D.J., McMurray, R.G., Murray, D.M., Almeida, M.J., Going, S., Norman, J.E. and Pate, R. 2004a. Defining accelerometer thresholds for activity intensities in adolescent girls. *Medicine and Science in Sports and Exercise*, 36 (7) pp.1259-1266.

Treuth, M.S., Sherwood, N.E., Baranowski, T., Butte, N.F., Jacobs, D.R.J., McClanahan, B., Gao, S., Rochon, J., Zhou, A., Robinson, T.N., Pruitt, L., Haskell, W. and Obarzanek, E.

2004b. Physical activity self-report and accelerometry measures from the Girls health Enrichment Multi-site Studies. *Preventive Medicine,* 38 (Suppl) pp.S43-S49.

Troiano, R.P., Berrigan, D., Dodd, K.W., Mâsse, L.C., Tilert, T. and McDowell, M. 2008. Physical Activity in the United States measured by accelerometer. *Medicine Science in Sports and Exercies,* 40 (1) pp.181-188.

Trost, S.G. 2001. Objective measurement of physical activity in youth: current issues, future directions. *Exercise and Sport Science Reviews,* 29 (1) pp.32-36.

Trost, S.G. 2007. State of the art reviews: measurement of physical activity in children and adolescents. *American Journal of Lifestyle Medicine,* 1 (4) pp.299-314.

Trost, S.G., Fees, B.S., Haar, S.J. and Murray, A.D.C.,L.K. 2012. Identification and validity of accelerometer cut-points for toddlers. *Obesity,* 20 (11) pp.2317-2319.

Trost, S.G., Loprinzi, P.D., Moore, R. and Pfeiffer, K.A. 2011. Comparison of accelerometer cut points for predicting activity intensity in youth. *Medicine and Science in Sports and Exercise,* 43 (7) pp.1360-1368.

Trost, S.G., McIver, K.L. and Pate, R.R. 2005. Conducting accelerometer-based activity assessments in field-based research. *Medicine and Science in Sports and Exercise,* 37 (Suppl 11) pp.S531-S543.

Trost, S.G., Pate, R.R., Freedson, P.S., Sallis, J.F. and Taylor, W.C. 2000. Using objective physical activity measures with youth: how many days of monitoring are needed? *Medicine and Science in Sports and Exercise,* 32 (2) pp.426-431.

Trost, S.G., Pate, R.R., Sallis, J.F., Freedson, P.S., Taylor, W.C., Dowda, M. and Sirard, J. 2002. Age and gender differences in objectively measured physical activity in youth. *Medicine and Science in Sports and Exercise,* 34 (2) pp.350-355.

Trost, S.G., Sirard, J.R., Dowda, M., Pfeiffer, K.A. and Pate, R.R. 2003. Physical activity in overweight and nonoverweight preschool children. *International Journal of Obesity,* 27 (7) pp.834-839.

Trost, S.G., Ward, D.S., Moorehead, S.M., Watson, P.D., Riner, W. and Burke, J.R. 1998. Validity of the computer science and applications (CSA) activity monitor in children. *Medicine and Science in Sports and Exercise,* 30 (4) pp.629-633.

Tryon, W.W. and Williams, R. 1996. Fully proportional actigraphy: a new instrument. *Behavior Research Methods, Instruments and Computers,* 28 (3) pp.392-403.

Tucker, P. 2008. The physical activity levels of preschool-aged children: a systematic review. *Early Childhood Research Quarterly,* 23 pp.547-558.

Tudor-Locke, C., Williams, J.E., Reis, J.P. and Pluto, D. 2004. Utility of pedometers for assessing physical activity: construct validity. *Sports Medicine,* 34 (5) pp.281-291.

Valanou, E.M., Bamia, C. and Trichopoulou, A. 2006. Methodology of physical-activity and energy-expenditure assessment: a review. *Journal of Public Health,* 14 pp.58-65.

Vale, S., Santos, R., Silva, P., Soares-Miranda, L. and Mota, J. 2009. Preschool children physical activity measurement: importance of epoch length choice. *Pediatric Exercise Science*, 21 (4) pp.413-420.

Vale, S., Silva, P., Santos, R., Soares-Miranda, L. and Mota, J. 2010. Compliance with physical activity guidelines in preschool children. *Journal of Sports Sciences*, 28 (6) pp.603-608.

Van Cauwenberghe, E., Jones, R.A., Hinkley, T., Crawford, D. and Okely, A.D. 2012. Patterns of physical activity and sedentary behaviour in preschool children. *International Journal of Behavioural Nutrition and Physical Activity* [online] 9 pp.138. Available at: http://www.ijbnpa.org/content/pdf/1479-5868-9-138.pdf [Accessed December 10 2012].

Van Cauwenberghe, E., Labarque, V., Trost, S.G., De Bourdeaudhuij, I. and Cardon, G. 2011. Calibration and comparison of accelerometer cut points in preschool children. *International Journal of Pediatric Obesity*, 6 (2-2) pp.e582-e589.

van Coevering, P., Harnack, L., Schmitz, K., Fulton, J.E., Galuska, D.A. and Gao, S. 2005. Feasibility of using accelerometers to measure physical activity in young adolescents. *Medicine and Science in Sports and Exercise*, 37 (5) pp.867-871.

Vanhees, L., Lefevre, J., Philippaerts, R., Martens, M., Huygens, W., Troosters, T. and Beunen, G. 2005. How to assess physical activity? How to assess physical fitness. *European Journal of Cardiovascular Prevention and Rehabilitation*, 12 (2) pp.102-114.

Vanhelst, J., Béghin, L., Rasoamanana, P., Theunynck, D., Meskini, T., Iliescu, C., Duhamel, A., Turck, D. and Gottrand, F. 2010a. Calibration of the RT3 accelerometer for various patterns of physical activity in children and adolescents. *Journal of Sports Sciences*, 28 (4) pp.381-387.

Vanhelst, J., Béghin, L., Turck, D. and Gottrand, F. 2010b. New validated thresholds for various intensities of physical activity in adolescents using the Actigraph accelerometer. *International Journal of Rehabilitation Research*, 34 (2) pp.175-177.

Vanhelst, J., Mikulovic, J., Bui-Xuan, G., Dieu, O., Blondeau, T., Fardy, P. and Béghin, L. 2012. Comparison of two Actigraph accelerometer generations in the assessment of physical activity in free living conditions. *BMC Research Notes* [online] (5), pp.187. Available at: http://www.biomedcentral.com/1756-0500/5/187 [Accessed May 20 2012].

Vanhelst, J., Theunynck, D., Gottrand, F. and Beghin, L. 2010. Reliability of the RT3 accelerometer for measurement of physical activity in adolescents. *Journal of Sports Sciences*, 28 (4) pp.375-379.

Vásquez, F., Salazar, G., Andrade, M., Vásquez, L. and Díaz, E. 2006. Energy balance and physical activity in obese children attending day-care centres. *European Journal of Clinical Nutrition*, 60 (9) pp.1115-1121.

Ventham, J.C. and Reilly, J.J. 1999. Reproducibility of resting metabolic rate measurement in children. *British Journal of Nutrition*, 81 (6) pp.435-437.

Verbestesl, V., van Cauwenberghe, E., De Coen, V., Maes, L., De Bourdeaudhuij, I. and Cardon, G. 2011. Within- and Between-Day Variability of Objectively Measured Physical Activity in Preschoolers. *Pediatric Exercise Science*, 23 pp.366-378.

Vincent, S. and Pangrazi, R.P. 2002. Does reactivity exist in children when measuring activity levels with pedometers? *Pediatric Exercise Science*, 14 (1) pp.56-63.

Ward, D.S., Evenson, K.R., Vaughn, A., Rodgers, A.B. and Troiano, R.P. 2005. Accelerometer use in physical activity: best practices and research recommendations. *Medicine and Science in Sports and Exercise*, 37 (Suppl 11) pp.S582-S588.

Welk, G.J. 2005. Principles of design and analyses for the calibration of accelerometry-based activity monitors. *Medicine and Science in Sports and Exercise,* 37 (Suppl 11) pp.S501-S511.

Welk, G.J., Blair, S.N., Wood, K., Jones, S. and Thompson, R.W. 2000a. A comparative evaluation of three accelerometry-based physical activity monitors. *Medicine and Science in Sports and Exercise,* 32 (Suppl 9) pp.S489-S497.

Welk, G.J., Corbin, C.B. and Dale, D. 2000b. Measurement issues in the assessment of physical activity in children. *Research Quarterly for Exercise and Sport*, 71 (Suppl 2) pp.S59-S73.

Wen, L.M., van der Ploeg, H.P., Kite, J., Cashmore, A. and Rissel, C. 2010. A validation study of assessing physical activity and sedentary behavior in children aged 3 to 5 years. *Pediatric Exercise Science,* 22 (3) pp.408-420.

Westerterp, K.R. 2001. Pattern and intensity of physical activity. *Nature,* 410 (6828) pp.539.

Westerterp, K.R. 2009. Assessment of physical activity: a critical appraisal. *European Journal of Applied Physiology,* 105 (6) pp.823-828.

Westerterp, K.R. and Kester, A.D. 2003. Physical activity in confined conditions as an indicator of free-living physical activity. *Obesity Research,* 11 (7) pp.865-868.

Wilkin, T.J., Mallam, K.M., Metcalf, B.S., Jeffery, A.N. and Voss, L.D. 2006. Variation in physical activity lies with the child, not his environment: evidence for an 'activitystat' in young children (EarlyBird 16). *International Journal of Obesity,* 30 (7) pp.1050-1055.

Williams, H.G., Pfeiffer, K.A., O'Neill, J.R., Dowda, M., McIver, K.L., Brown, W.H. and Pate, R.R. 2008. Motor skill performance and physical activity in preschool children. *Obesity,* 16 (6) pp.1421-1426.

Winkler, E.A., Gardiner, P.A., Healy, G.N., Clark, B.K. and Sugiyama, T. 2009. Distinguishing true sedentary from accelerometer non-wearing time: accuracy of two automated wear-time estimations. *Medicine and Science in Sports and Exercise,* 41 (5 Suppl 1) pp.171-172.

Yamamoto, S., Becker, S., Fischer, J. and De Brock, F. 2011. Sex differences in the variables associated with objectively measured moderate-to-vigorous physical activity in preschoolers. *Preventive Medicine,* 52 pp.126-129.

Yang, C.C. and Hsu, Y.L. 2010. A review of accelerometry-based wearable motion dectectors for physical activity monitoring. *Sensors*, 10 (8) pp.7772-7788.

Zhang, K., Pi-Sunyer, F.X. and Boozer, C.N. 2004. Improving energy expenditure estimation for physical activity. *Medicine and Science in Sports and Exercise,* 36 (5) pp.883-889.

Zhang, K., Werner, P., Sun, M., Pi-Sunyer, F.X. and Boozer, C.N. 2003. Measurement of human daily physical activity. *Obesity Research,* 11 (1) pp.33-40.

APPENDICES

APPENDIX I: Summary of studies of physical activity and sedentary behaviour of preschool children

Appendix Table I.i: Summary of studies into preschool children's physical activity and sedentary behaviour

Measurement approach Author	**Country**	**Measurement Protocol**	**No. participants (n), gender, age range, mean (SD)**	**Key findings** **Unless stated figures are the mean (SD)**	**≥ 60 min/day MVPA**	**≥180 min/day TPA**	**≤60 min/day or ≤120 min/day of ST**
Accelerometer (Actigraph)							
Alhassan et al. (2007)	USA	2 days at baseline (week day only)	32, 12 girls, 20 boys 3.6 (0.5) y CON: control (n=17) RECESS: intervention (n=15)	Percentage time in: **Sed**: CON: 93.9 (3.0)%, RECESS: 94.9 (1.9)% **LPA**: CON:4.1 (1.7)%, RECESS: 3.7 (1.3)% **MVPA**: CON: 2.0 (1.6)%, RECESS 1.4 (0.9)%	No	No	NA
Beets et al. (2011a)	USA	Data collected over 2 weeks; (inclusion ≥ 1 day)	379, 52.3% girls, 47.7% boys; 3 - 5 y	**MVPA**: ranged from: 39.5 (22.5) min/day to 269.0 (70.8) min/day **TPA**:127.3(47.5) min/day to 402.5 (83.6) min/day No difference between weekend & week	Unclear – depends on data processing	Unclear – depends on data processing	NA
Burdette et al. (2004)	USA	3 days	250, 107 girls, 143 boys, 3.7 y	Mean outdoor play: 146 min/day	NA	NA	NA
Bürgi et al. (2010)	Switzerland	3 days (including 2 week, 1 weekend day) ≥ 6 h Compared German speaking (GS) & French speaking (FS) children	524 278 girls, 278 boys, 5.1 (0.60) y	**MVPA**: GS: 400 (110); FS: 361 (101) number of 15-s intervals/day ($p < 0.001$) **TPA**: GS: 771(169); FS: 684(151) number of 15s intervals/day ($p < 0.001$) **Sed**: GS 1276 (216); FS: 1400 (253) number of 15-s intervals/day ($p < 0.001$) **TV time**: GS: 45 (43); FS: 67 (50) min/day ($p = 0.001$)	NA	NA	No: FS > 60 min/day but < 120 min/day Yes GS

Measurement approach Author	Country	Measurement Protocol	No. participants (n), gender, age range, mean (SD)	Key findings Unless stated figures are the mean (SD)	≥ 60 min/day MVPA	≥180 min/day TPA	≤60 min/day or ≤120 min/day of ST
Cardon and Bourdeaudhij (2008)	Belgium	4 days (including 2 week days, 2 weekend days) Mean monitoring time 11.8 (1.5) h/day	76, 39 girls, 37 boys, 4 - 5 y	**Sed**: 85% of time, 598 (78) min/day **MVPA**: 5% time, 34 (27) min/day 7% achieved ≥ 60 min/day of MVPA 26% achieved ≥ 120 min/day of total activity No difference between 4 y olds & 5y olds activity Higher counts on weekend & more sedentary on week(p < 0.01) No gender differences for total activity, boys spent more time in MPA (p < 0.01)	No	NA	NA
Cliff et al. (2009a)	Australia	≥ 3 days ≥ 6 h	46, 21 girls, 25 boys 4.3 (0.7) y	**MVPA**: 23 min/day No difference between girls & boys in MVPA Object control skills were positively associated with PA in boys Locomotor skills were negatively associated with PA in girls	No	No	NA
Cliff et al. (2011)	Australia	≥ 4 days ≥10 h/day Screen behaviour- parental report Movement skills proficiency RCT – subjects assigned to a physical activity group (PA)or diet group (D) or PA & diet group (PAD)	137 at baseline 5.5 - 9 y Obese or Overweight children.	**MVPA**: 194 min/day **ST**: 171 min/day **MVPA**: PA group: 23.8(6.9)%, Diet: 24.7(8.0)%; Activity & diet: 28.1(6.8)% **TV or DVD viewing**: PA: 885 min/week Diet:1000 min/week; 960 min/week (~126.4 - 142.9 min/day)	Yes	NA	No≥ 60 min/day

Measurement approach Author	Country	Measurement Protocol	No. participants (n), gender, age range, mean (SD)	Key findings Unless stated figures are the mean (SD)	≥ 60 min/day MVPA	≥180 min/day TPA	≤60 min/day or ≤120 min/day of ST
Eijkemans et al. (2008)	Netherlands	≥ 3 days (including 1 weekend day) ≥ 400 min/day	305, 153 girls, 152 boys, 4 - 5y	Girls: 606 cpm, **VPA**: 5.7 min/day, **MVPA**: 2.6 episodes ≥ 1min Boys: 650 cpm, **VPA**:5.4 min/day, **MVPA** 2.6 episodes ≥ 1min Underweight: (n = 35) 609 cpm, **VPA**: 4.8 min/day, **MVPA**: 2.8 episodes ≥ 1min Normal weight: (n = 242): 629 cpm, **VPA**: 5.7 min/day, **MVPA**: 2.8 episodes ≥ 1min Overweight (n = 22): 655cpm, VPA: 5.7 min/day, **MVPA**: 3.0 episodes ≥ 1min Obese (n = 6): 576cpm, **VPA**: 3.6 min/day, **MVPA**: 1.2 episodes ≥ 1min Difference obese & normal weight in **VPA** ($p <$ 0.05) & difference in episodes **MVPA**	NA	NA	NA
Fisher et al. (2005a)	Scotland	6 days	394, 185 girls, 209 boys, 3 & 5 y, 4.2 (0.5) y	**Sed**: 76.3%, **LPA**:20.3%; **MVPA**: 3.4% of total time; 769 (192) cpm **TPA** & **MVPA** significantly higher in boys ($p <$ 0.001) **Sed** significantly lower in boys ($p < 0.001$)	No	NA	NA
Fisher et al. (2005b)	Scotland	Younger children 3 day (including 2 week days & 1 weekend day). ≥ 6 h Older children 7 consecutive day. ≥ 6 h	209, 108 girls,101 boys, 3 - 5 y, 4.8 y	**TPA**, **MVPA** & **LPA** significantly lower in spring than summer **Sed** significantly higher in spring than summer. Differences small so possibly not biologically meaningful. Spring: **Sed**: 79.5%, **LPA**:17.0%; **MVPA**:2.7% Summer: **Sed**: 74.2%; **LPA**: 21.7%; **MVPA**:3.8% Fall: **Sed**: 76.1%; **LPA**:20.1%; **MVPA**:4.1% Winter: **Sed**: 76.6%; **LPA**: 19.5%; **MVPA**:2.1%	No	NA	NA

Measurement approach Author	**Country**	**Measurement Protocol**	**No. participants (n), gender, age range, mean (SD)**	**Key findings Unless stated figures are the mean (SD)**	**≥ 60 min/day MVPA**	**≥180 min/day TPA**	**≤60 min/day or ≤120 min/day of ST**
Gabel et al. (2011)	Canada	≥ 5 h on ≥ 4 day (including 1 weekend day) Mean wear time, baseline: 10.6 (0.9) h; follow up: 11.6 (0.9)	17, 4.4 (0.8)y	Baseline: **TPA**: 34.4 (4.7)%; **MVPA**: 11.8 (1.7)%; **LPA**: 22.6 (3.2)% Follow up: **TPA**: 36.8 (5.2)%; **MVPA**: 13.5 (2.9)%; **LPA**: 23.3 (3.0)%	Yes	Yes	NA
Heelan & Eisenmann (2006)	USA	≥ 4 day (including 1 weekend day) ≥ 480 mins (8 h).	100, 52 girls, 48 boys 4 - 7 y	**MPA**:241.5 (48.8) min/day; **VPA**:32.3 (17.1) min/day; **MVPA**:273.8 (59.1) min/day. No significant differences between gender	Yes	Yes	NA
Hinkley et al. (2012c)	Australia	≥ 3 day (including 1 weekend day) (mean = 6.9 day). ≥ 50% of usual wake time on each day (mean: 647.5 min/day).	703, 315 girls, 388 boys 4.5 y	**LPA**:11.7(2.4)%; **MPA**:3.4(1.9)%; **VPA** 1.4 (0.9)%; Total activity:16.4(4.2)%. 5.1% meet recommendation of ≥ 3h/day of total physical activity 58.9% meet recommendation of ≤ 2 h/day of screen time **ST**:112.5(73.2) min/day Boys spent significantly more time in **LPA**, **MPA** & in total activity than girls (p < 0.001)	No	No	Yes ≤ 120 min/day
Hinkley et al. (2012a)	Australia	≥ 4 day (including ≥ 3 week day & 1 weekend day)	939, 45% girls, 55% boys, 4.5 y	Boys higher percentage of daily time in **TPA** than girls across entire week, week & weekend day (p < 0.001). Boys & girls spent greater mean percentage time in activity on weekend than week. Boys (p < 0.001) girls (p = 0.004). **TPA** declined with age for both girls & boys.	NA	NA	NA

Measurement approach Author	Country	Measurement Protocol	No. participants (n), gender, age range, mean (SD)	Key findings Unless stated figures are the mean (SD)	≥ 60 min/day MVPA	≥180 min/day TPA	≤60 min/day or ≤120 min/day of ST
Hume et al. (2012)	Australia	≥ 4 day (including ≥ 1 weekend day) Valid day10 - 14 h of wear time	373, 206 girls, 167 boys, 5 - 12 y Girls 9.6 (2.0) y; Boys 9.2(2.1) y.	Girls **MVPA**: 156.4 (62.9) min/day Boys **MVPA**:195.0(68.7) min/day; ($p < 0.001$) **ST**: boys: 160.0 (97.1); girls:141.9 (83.7) boys: 89.2% **MVPA**≥ 60 min/day girls: 79.6% **MVPA** ≥ 60 min/day boys: 37.7% **ST** ≤120 min/day girls: 47.1% **ST** ≤120 min/day	Yes	NA	No ≥ min/day
Jackson et al. (2003)	Scotland	≥ 3 day (including 1 weekend day) Follow up at 1 year	Baseline: 104, 52 girls, 52 boys , 3 y Follow up: 60, 30 girls, 30 boys at 4 y	Baseline: 669 (165) cpm Follow up: 849 (252) cpm Activity increased over 1 year ($p < 0.001$) Boys more active than girls ($p < 0.001$)	NA	NA	NA
Janz et al. (2001)	USA	≥ 4 day (including 1 weekend day)	368, 189 girls, 179 boys, 4 - 6 y	Boys more **TPA** & **VPA** than girls ($p < 0.05$)	NA	NA	NA
Janz et al. (2002)	USA	≥ 4 day (including 1 weekend day)	434, 231 girls, 203 boys, 4 - 6 y	Boys **MVPA**:277.1 (50.8) min/day Girls **MVPA**: 263.0(48.2) min/day Boys more **MVPA** & **TPA** than girls ($p < 0.001$)	Yes	Yes	NA
Janz et al. (2004)	USA	≥ 4 day, (including 1 weekend day) Valid day ≥ 8 h/day for 3 day	436, 232 girls, 204 boys; 4 - 7 y	Boys 244 (43), 267 (44), 38 (19) min/day in **Sed**, **MPA** & **LPA** respectively Girls 251(48), 262 (44), 28(14) min/day in **Sed**, **MPA** & **LPA** respectively Boys more active than girls ($p < 0.05$)	Yes	Yes	NA
Janz et al. (2010)	USA	≥ 4 day, (including 1 weekend day) for 5 & 8 y olds ≥ 8 h/day for ≥ 3 day	333, 185 girls, 148 boys, 5, 8 & 11 y olds	5 y olds: Girls **MVPA**: 24.2 (13.4) min/day Boys **MVPA**: 31.1 (16.3) min/day Boys more active than girls ($p < 0.05$)	No	NA	NA

Measurement approach Author	**Country**	**Measurement Protocol**	**No. participants (n), gender, age range, mean (SD)**	**Key findings Unless stated figures are the mean (SD)**	**≥ 60 min/day MVPA**	**≥180 min/day TPA**	**≤60 min/day or ≤120 min/day of ST**
Kelly et al. (2005)	Scotland	7 day for ≥ 6 h	41, 21 girls, 20 boys, 4.3 - 6 y	**TPA**: 726 cpm; **Sed**: 78%, **LPA**:19%; **MVPA**:3%; boys had significantly higher total activity than girls (p = 0.0015); more time in **MVPA** & less time in **Sed**	No	NA	NA
Kelly et al. (2006)	Scotland	Study 1 6 day ≥ 6 h Study 2 7 day ≥ 6 h	Study 1: 339, 4.2 (0.3) y Study 2: 78 5.6 (0.4) y	Study 1 & 2: **MVPA**: 3% of time **Sed**: 77% & 78% of time	No	NA	NA
Kelly et al. (2007)	Scotland	Longitudinal study 24 month follow-up 3 day (including 1 weekend day); 5 day follow up ≥ 6 h (median = 10.2 h)	42, 21 girls, 21 boys 3 - 5 y Baseline: 3.8 (0.5) y Follow-up: 5.5 (0.4) y	Baseline: **Sed**:80.5%; **MVPA**: 2.2% Follow-up: **Sed**:74.1%; **MVPA**:4.1%	No	NA	NA
Martinez-Gomez et al. (2009)	Spain	≥ 3 week day (including 1 weekend day)	110, 56 boys, 54 girls, 3 - 8 y 97% Caucasian	**MVPA**: 215 min/day No significant difference in **MVPA** between boys & girls	Yes	Yes	NA
Metallinos-Katsaras(2007)	USA	7 d (mean; 6.6 day),	56, 2 - 5 y 38% overweight or at risk of overweight	**LPA**:416.2 (75.9) min/day (60.8%); **MPA**:243.7(50.1) min/day (35.3%); **VPA**:24.5 (13.9) min/day; **VVPA** 3.9(3.4) min/day (4.0% in **VPA** or **VVPA**); **TPA**: 272.2 (60.1) min/day Boys more active than girls (p < 0.05) higher daily vigorous activity (p < 0.05)	Yes	Yes	NA
Metcalf et al. (2002)	England	7 consecutive day	82, 45 girls, 37 boys, 4.8 y	Less **TPA** on school day than on weekends (p = 0.0006) Boys were more active than girls (p = 0.04)	NA	NA	NA

Measurement approach Author	Country	Measurement Protocol	No. participants (n), gender, age range, mean (SD)	Key findings Unless stated figures are the mean (SD)	≥ 60 min/day MVPA	≥180 min/day TPA	≤60 min/day or ≤120 min/day of ST
Montgomery et al. (2004)	Scotland	Accelerometry & DLW ≥ 3 day for preschool age (n = 36) 7 - 10 day for first school year (n = 78) 5 - 10 h/day	104 52 girls, 52 boys, 5.3 y	**TPA** boys: 848; girls 719 counts/min **Sed**: boys: 73%; girls:79% of total time **LPA**: boys: 23%; girls: 18% of total time **MVPA**: boys: 4%; girls: 3% of total time Boys more total counts ($p < 0.001$) less **Sed** ($p < 0.0002$); more LPA ($p < 0.001$); more **MVPA** ($p < 0.0068$) Boys had significantly higher **TEE** ($p < 0.0003$), **AEE** ($p < 0.0002$) & **PAL** ($p < 0.0001$)	No	NA	NA
Obeid et al. (2011)	Canada	Accelerometry 7 day	30 3 - 5 y	**TPA**: 219.7 (31.9) min/day **MVPA** 75 min/day	Yes	Yes	NA
O'Dwyer et al. (2011)	England	Accelerometry 7 day	50, 23 girls, 27 boys 4.4 (0.5)y Overweight (OW) (n = 17) Non overweight (NOW) (n = 32)	**MVPA**: boys OW week: 38.6(9.1) min/day; weekend: 34.0 (11.9) min/day; NOW week: 45.2 (20.3) min/day; weekend: 38.0 (10.4) min/day **MVPA**: girls o OW t week: 38.0 (10.5) min/day; weekend: 28.9 (9.5) min/day; NOW week: 43.3 (17.0) min/day; weekend: 42.4 (26.4) min/day. **Sed**: boys OW week: 652.6(168.6) min/day, weekend: 863.7(164.4) min/day; NOW week: 751.3(146.7) min/day; weekend: 684.0(198.1) min/day. **Sed**: girls OW week: 668.0(150.8) min/day; weekend: 673.0(200.4) min/day; NOW week: 672.4(117.4) min/day; weekend: 757.0 (203.0) min/day. % > 60 min MVPA per day**:** boys OW: 0%; boys NOW: week: 25%; weekend: 20%; girls: OW: 0%; NOW week: 15% weekend: 20%	No	NA	NA

Measurement approach Author	Country	Measurement Protocol	No. participants (n), gender, age range, mean (SD)	Key findings Unless stated figures are the mean (SD)	≥ 60 min/day MVPA	≥180 min/day TPA	≤60 min/day or ≤120 min/day of ST
Pfeiffer et al. (2009)	USA	8 - 10 day (including 1 weekend day) Up to 5 week day & 2 weekend day used for analysis Days < 5 h & >18 h excluded. < 3 day of data excluded	331, 163 girls, 168 boys; 3 - 5 y 4.3 (0.6) y	**Sed**: 7.6(2.1) min/h, **Non Sed (LPA+ MPA+ VPA):** 27.2 (3.9) min/h	NA	NA	NA
Quigg et al. (2010)	New Zealand	6 day	176, 94 girls, 82 boys, 5 - 10 y 7.6 y	Mean total daily activity counts Girls: 381937 (363 007); boys: 381937 (363 007) 2% of children's weekly activity was located in a city park with a playground	NA	NA	NA
Reilly et al. (2004)	Scotland	2 - 3 day at 3 yr; 7 day at 5 yr	150, 73 girls, 77 boys, 3 & 5 y	3 y **Sed**: 79% **MVPA**: 2% 5 y **Sed**: 76%; **MVPA** 4%	No	NA	NA
Reilly et al. (2006b)	Scotland	6 day	545, 272 girls, 273 boys, 4.2 (0.2) y	At baseline: Control group (n = 277) **Sed**: 66.9% **MVPA**: 3.0% time Intervention group (n = 268) **Sed**: 69.3%, **MVPA**: 2.6% time	No	NA	NA
Sarzynski et al. (2010)	USA	≥ 4 day ≥ 8 h	132, 64 girls, 68 boys 3 - 12 y 7.1 (1.9) y	**MVPA**: 63.1 (30.9) min/day **MVPA** girls: 55.0 (28.4); boys: 70.8 (31.4) No significant difference between time in **MVPA** between genders	Yes	NA	NA
Spittaels et al. (2012)	Belgium	6 day (including 2 weekend day)	207 93 girls, 114 boys 4.5 (0.8) y	**Sed**: 51(7)% of time; **LPA**: 41 (5)% of time; **MVPA**: 8 (3)% of time; **TPA**: 566 (133) cpm	No	Unclear	NA

Measurement approach Author	Country	Measurement Protocol	No. participants (n), gender, age range, mean (SD)	Key findings Unless stated figures are the mean (SD)	≥ 60 min/day MVPA	≥180 min/day TPA	≤60 min/day or ≤120 min/day of ST
Telford et al. (2005)	Australia	4 day (including 1 weekend day) days < 10,000 counts excluded VPA > 6 h excluded (Accelerometry & CLASS questionnaire)	237, 49% girls, 51% boys, 5 - 6 y 887 46% girls, 54% boys, 10 - 12 y	Boys: (5 - 6y): **Sed**:284 (109.1); **LPA**:254(38.0); **MPA**:226.6 (51.5); **VPA**:43.4 (18.9) min/day Girls: (5 - 6y) **Sed**:297(102.6); **LPA**:254.5(38); **MPA**:214.1 (43.6); **VPA**:32.8(13.7) min/day Younger children (5 - 6y) 52% of time **Sed** Significant gender differences in **MPA** (p < 0.05); 99% of children ≥ 60 min/day in **MVPA**	Yes	Yes	NA
Trost et al. (2003)	USA	1 - 11 day (DO for 1 h on 3 day)	245, 127 girls, 118 boys, 3 - 5 y	**MVPA**: non-overweight girls: 41.6 (12.5)% of time & non-overweight boys: 47.6 (12.7)% of time **MVPA**: overweight girls:42.2 (12.8)% of time & overweight boys: 39.0 (12.5)% of time.	Yes possibly	NA	NA
Vale et al. (2010)	Portugal	≥ 4 day (including 1 weekend day) ≥ 10 h valid day	245, 105 girls, 140 boys, 3.5 - 6.0 y	**TPA**: week:143.8 (43.3) min/day; weekend: 123.9 (41.8) min/day; **MVPA**: week: 102.3 (31.2) min/day; weekend: (33.7) min/day Boys significantly more activity (p < 0.05). Significantly more time active & in MVPA at weekend. **TPA**:17.1% weekend,16.1% week **MVPA**: 12.1% weekend, 11.5% week	Yes	No	NA
Van Cauwenberghe et al. (2012)	Australia	≥ 4 day (including 3 week day & 1 weekend day)	703, 316 girls, 387 boys, 4.6 (0.7) y	**Sed**: girls week: 309 (70) min/day, boys: 306 (77) min/day girls: weekend: 300 (78) min/day , boys: 293 (23) min/day **MVPA**: girls week: 27 (17) min/day, boys: 32 (21) min/day ; girls weekend: 30 (21) min/day; boys: 36 (25) min/day	No	NA	NA

Measurement approach Author	Country	Measurement Protocol	No. participants (n), gender, age range, mean (SD)	Key findings Unless stated figures are the mean (SD)	≥ 60 min/day MVPA	≥180 min/day TPA	≤60 min/day or ≤120 min/day of ST
Verbestel et al. (2011)	Belgium	≥ 3day (including 1 weekend day)	213, 48% girls, 52% boys, 4.98 (0.88) y	**TPA**: 586.42 (147.36) cpm **Girls**: 569.91 (120.16) cpm; boys: 603.39 (166.75) cpm (p = 0.094) Week: 597.26 (149.27) cpm; weekend: 579.92 (216.16) cpm (p = 0.10)	NA	NA	NA
Wen et al. (2010)	Australia	≥ 4 day ≥ 8h	31, 12 girls, 19 boys 3 - 5 y 3.5y	**MVPA**: 34 (17.0) min/day; **Sed**: 630 (124) min/day **ST** (parental report): 69 min/day	No	NA	No>60 min/day
Wilkin et al. (2006)	England	≥ 5 day (including 4 week day & 1 weekend day) 1 min epoch	272, 120 girls, 152 boys, 4.9 y 247, 114 girls, 133 boys, 5.9 y	**TPA**: 4.9y: 37.5 (7.7); 5.9 y: 37.4 (7.7) (units x 10^5 /w) High intensity: 4.9 y: 12.2 (5.5); 5.9 y: 12.7 (6.1) No change in PA year to year Girls less active than boys at 4.9 y ($p < 0.001$) & 5.9 y ($p < 0.02$), girls less high intensity than boys at 4.9 y ($p < 0.001$) & 5.9y ($p < 0.01$) TV viewing in young children did not impact on PA	NA	NA	NA
Williams et al. (2008)	USA	7164 ≥ 3 day - 7 day Up to 5 week day & 2 weekend day <5 h or >18 h excluded (mean time 12.7 (1.6) h)	198, 49.5 % girls, 50.5% boys, 3 - 4 y	**Sed**: 54.8 (6.3); **LPA**: 32.6 (4.3); **MVPA**: 12.6 (3.6); **VPA**: 4.5 (1.0) % of time. **Sed**: ~ 7 h/day **MVPA**: ~ 90 min/day No difference between 3 & 4 y olds	Yes	NA	NA

Measurement approach Author	Country	Measurement Protocol	No. participants (n), gender, age range, mean (SD)	Key findings Unless stated figures are the mean (SD)	≥ 60 min/day MVPA	≥180 min/day TPA	≤60 min/day or ≤120 min/day of ST
Yamamoto et al. (2011)	Germany	≥ 2 day (including 1 week & 1 weekend day) Mean: (3.8 day Mean weekend: 13.4 h/day week: 12.8 h/day	979, 49% girls, 51% boys, 3 - 6 y	Girls: negative association between PA & age, suggesting younger girls more likely to engage in PA, PA had positive association with health; positive association with higher TV viewing. Positive association with higher parental PA. Boys: PA positively associated with desire to be active	NA	NA	NA
Accelerometer (Actical)							
Dolinsky et al.(2011)	USA	≥ 3 day (including 2 week day & 1 weekend day)	337 142 girls, 195 boys, 3.5 (1.1) y	**Sed**: 6.1 h/day **MVPA**: 14.9 (9.5) min/day	No	NA	NA
Taylor et al. (2009)	New Zealand	Actical & parental report-questionnaire developed for study. Longitudinal study, baseline & annual follow up for 2 y 5 consecutive day of 24 h accelerometry data collection	244, 44% female, 56% male, 3, 4 & 5 y	**TPA**: 3 y:81 (54) min/day; 4 y:72 (53) min/day; 5y: 57 (39) min/day **Sed**: 3 y:179 (92) min/day; 4y:178 (91) min/day; 5 y:160 (74) min/day **MVPA** 3 y:42 min/day; 4 y:16.5 min/day; 5 y: 22 min/day Declines with total PA (3 - 4 y, p = 0.001 & 4 - 5 y, p = 0.054) & **MVPA** with age. No difference between weekend day & weekday gender; or seasonality (p > 0.05); **ST**=90 min/day Other sed activities = 70 - 90 min/day	No	No	Yes: ≤ 120 min/day No: ≥ 60 min/day
Accelerometer (Caltrac)							
Moore et al. (1995)	USA	Longitudinal study 6 months apart 5 day	97, 39 girls, 58 boys 4.0 y at baseline	10.9 (1.9) count/h; boys 11.3 (2.0); girls 10.3 (1.7) min/h;	NA	NA	NA

Measurement approach Author	Country	Measurement Protocol	No. participants (n), gender, age range, mean (SD)	Key findings Unless stated figures are the mean (SD)	≥ 60 min/day MVPA	≥180 min/day TPA	≤60 min/day or ≤120 min/day of ST
Sigmund et al. (2009)	Czech Republic	7 day Yamax SW-200 pedometer	176, 84 girls, 92 boys Kindergarten:5.7 (0.5) y Elementary school: 6.7 (0.5) y	Week: Mean AEE girls: 11.5 $kcal \cdot kg \cdot d^{-1}$, boys: 12.9 $kcal \cdot kg \cdot d^{-1}$, Steps/day girls: 9923, boys: 11864. Weekend: Mean AEE: girls:11.5 $kcal \cdot kg \cdot d^{-1}$, boys: 12.7 $kcal \cdot kg \cdot d^{-1}$, Steps/day girls: 10606, boys: 11182. No difference weekend/week	10,000 steps – yes	NA	NA
Accelerometer (RT3)							
Burdette et al. (2005)	USA	≥ 3 day (including 1 weekend day)	250, 107 girls, 143 boys 29 - 52 month, 44 months	146 (113) min/day outdoor playtime – recall. Seasonal difference in proxy report not in accelerometry measures. No difference in outdoor play measures-recall 667 (186) vector magnitude per minute Boys higher accelerometry measures ($p < 0.01$) 113 (79) min/day watching TV or video	NA	NA	No ≤ 60 min/day Yes ≤120 min/day
Vásquez et al. (2006)	Chile	≥ 3day (including 1 weekend day)	19, 9 girls, 10 boys, 3 - 5 y	**Sed/LPA**: weekend: 52%; week:54%, day care: 62%. **MPA** weekend: 3%, week: 4%	NA	NA	NA
Accelerometer (Actiwatch)							
Butte et al. (2007)	USA	3 day 24 h monitoring VO_{2peak} ramp protocol HR>195 bpm or RQ>1.0 maximal effort	897, 4 - 19 y 10.8 (3.8) y	Mean counts: 1410 (35) min/day Counts decreased with age (p = 0.001). Total counts higher in boys (p = 0.001) & non overweight children (p = 0.002) Fitness: VO_{2peak} lower in overweight children & girls (p = 0.001), difference between overweight & non overweight boys	NA	NA	NA

Measurement approach Author	Country	Measurement Protocol	No. participants (n), gender, age range, mean (SD)	Key findings Unless stated figures are the mean (SD)	≥ 60 min/day MVPA	≥180 min/day TPA	≤60 min/day or ≤120 min/day of ST
Chen et al. (2002)	Japan	3 day Actiwatch Calorie counter (measures step count, energy expenditure from PA, AEE, TEE, BMR.	21, 9 girls, 12 boys, 3 - 4 y	Mean activity counts: 404.5 (155.08) Boys & girls (p = 0.118). Steps: 12638 (2849) steps/day (p = 0.691) Gender difference for **TEE** (p < 0.01); No gender difference for **AEE** (p = 0.646)	NA	NA	NA
Finn and Ulmann (2004)	USA	3 week day 2 childcare centres VPA>250 counts/15 s	28, 16 girls, 12 boys, 3 - 5 y	VPA: cold months: 67.23 (21.65) min/day; warm months: 66.48 (25.29) min/day. No difference between different temperatures (cold & warm seasons).	Yes	NA	NA
Finn, Johannsen, Specker (2002)	USA	2 day (48 h) ≥ 24 h of data Actiwatch (Mini Mitter Co-biaxial CARS Validation with CARS with subgroup (n = 40)	214, 108 girls, 106 boys, 3 - 5 y	Mean daily counts (x10,000): girls: 26.3 (0.7); boys: 28.5 (0.8) (p = 0.03) counts/day. Daily counts between 9am - 5pm (%) girls: 53 (1.1)%, boys: 53.3 (1.1). Percentage time spent in vigorous activity: girls: 4.5 (0.2)%, boys: 5.2 (0.2) % (p = 0.02) Children born pre-term lower activity counts & less time in vigorous activity.	NA	NA	NA
Firrincieli et al. (2005)	USA	6 - 7 day	54, 33 girls, 21 boys, 3 - 5 y	10 min bouts of PA: asthmatic children: 1041 non asthmatic children: 1610, normalised over 7 day.	NA	NA	NA
Jackson et al. (2009)	Scotland	7 day Doubly labeled water	89, 42 girls, 47 boys 2 - 6 y	442.4 (102.6) cpm No difference between boys & girls Negative association between TV viewing & **PA** 1.87 (0.99) h/day TV viewing TV viewing associated with greater levels of body fat **TEE** girls: 5850 (1523) boys: 6756 (1371) (p < 0.05)	NA	NA	No ≤60 mins Yes ≤120 mins

Measurement approach Author	Country	Measurement Protocol	No. participants (n), gender, age range, mean (SD)	Key findings Unless stated figures are the mean (SD)	≥ 60 min/day MVPA	≥180 min/day TPA	≤60 min/day or ≤120 min/day of ST
Specker and Binkley (2003)	USA	48 h at 0, 6 & 12 m	178, 84 girls, 94 boys 3 - 4 y	Baseline: **MPA**: 12.1 - 14.0%, **VPA**: 4.5 - 5.4% of time. Av daily counts: 259 000 - 297 000	NA	NA	NA
Accelerometer (ActivTracer & Lifecorder)							
Tanaka & Tanaka (2009)	Japan	≥ 3 day (including 1 weekend day) analysed Uniaxial & triaxial accelerometers	157, 69 girls, 88 boys 4 - 6 y 5.9 (0.5) y	ActivTracer (triaxial):MVPA: 102.0 (32.0) min/day; girls 88.8 (28.9); boys: 122.3 (30.7) min/day Difference between boys & girls (p < 0.05) MVPA ≥ 60 min/day = 92.4% of children Lifecorder (uniaxial): step counts: 13037 (2846), girls: 12255 (2823); boys: 13650 (2726). Difference between boys & girls (p < 0.05)	Yes	NA	NA
Accelerometer & Pedometers							
Cardon and De Bourdeaudhuij (2007)	Belgium	≥ 4 day (including 2 weekend day) Actigraph accelerometer Digiwalker pedometer & diaries	122, 63 girls, 59 boys 4 - 5.9 y Actigraph with 76, 39 girls, 37 boys 4 - 5.9 y	9980 step counts/day; 18.68 steps/min 60 min **MVPA**=13,874 step counts 8% (n = 10) achieved ≥ 60 min **MVPA** per day Week activity higher than weekend (p < 0.001)	No	NA	NA
DO (OSRAP)							
Dowda et al. (2004)	USA	1 h 2 - 3 day	266, 140 girls, 126 boys 3-5 y	**MVPA** 27% of observed time	NA	NA	NA
Bower et al. (2008)	USA	10,240 observations made over 2 day	20 childcare centres 3 - 5 y	**Sed** (CARS 1 - 2): 55% ; **MVPA**: (CARS 4 - 5): 12% 4 - 7% of time in **MVPA**.	No likely	NA	NA
DO (CARS)							

Measurement approach Author	Country	Measurement Protocol	No. participants (n), gender, age range, mean (SD)	Key findings Unless stated figures are the mean (SD)	≥ 60 min/day MVPA	≥180 min/day TPA	≤60 min/day or ≤120 min/day of ST
Baranowski et al. (1993)	USA	5 - 6 h on ≥ 4 day	191, 101 girls, 90 boys 3 - 4 y	Overall mean PA was low, Outdoor play associated with PA; gender; month & location Boys more active than girls	NA	NA	NA
DuRant et al. (1994)	USA	≥ 4 day, 6 - 12 h/day	191, 101 girls, 90 boys, 3 - 4 y	Mean PA highest during outdoor play, lowest during TV watching.	NA	NA	NA
Pedometry							
Kambas et al. (2012)	Greece	7 consecutive day Omron walking style pro HJ-720IT-E2	232, 114 girls, 118 boys 5 - 5.5y 5.4 y	Steps/day: 7676 (1893) Aerobic walk time: 12.8 (17.5) Aerobic steps/day: 1486 (1995)	NA	NA	NA
Duncan et al. (2008)	New Zealand	≥ 5 day (including 3 week day & 2 weekend day) Multiday memory (MDM) pedometer NL-2000 Daily step count ≥ 30,000 or ≤ 1000 were removed	1115, 579 girls, 536 boys, 5 - 12 y	Mean step counts week: Boys: 16,100; girls: 14,200 Mean step counts weekend: Boys: 12,900, girls: 11,300 Mean temperature +ve effect larger at weekends (26% more steps) for boys & moderate for week (11% more steps), trivial effect at weekends in girls , small effect on week (16% more steps)	Yes > 10,000 steps/day	NA	NA
Questionnaires /proxy report							
Carson et al. (2010)	Canada	Children's Leisure Activities Study Survey (CLASS)	Winter: n = 260, Spring: n = 507, Summer: n = 684, Fall: n = 264, 48.9% girls, 51.1 % boys 4 - 5 y	**Active play**: Winter: 180 min/week; Spring: 218 min/week; Summer: 270 min./week, Fall: 197.5 min/week **TPA**: Winter: 494.0 min/week; Spring: 540 min/week; Summer: 600 min/week; Fall: 535 min/week Difference in seasons for **TPA** ($p < 0.001$) Active play ($p < 0.0001$), week ($p < 0.0001$); weekend ($p < 0.006$)	NA	NA	No

Measurement approach Author	Country	Measurement Protocol	No. participants (n), gender, age range, mean (SD)	Key findings Unless stated figures are the mean (SD)	≥ 60 min/day MVPA	≥180 min/day TPA	≤60 min/day or ≤120 min/day of ST
Okley et al. (2009)	Australia	Physical activity & Exercise Questionnaire PAEC-Q	266, 126 girls, 140 boys, 3.96 (0.76) y	**TPA** week: 55.7% spent ≥ 3 h/day **TPA** weekend: 9.0 (3.3)% spent ≥ 3 h/day **TV viewing:** week: 3.0 (3.7)% ≤ 2 h/day **TV viewing:** weekend**:** 70.4 (4.0)% ≤ 2 h/day	NA	Yes	Yes
Cox et al. (2012)	Australia	Eating & Physical Activity Questionnaire Reported TV viewing	135 81 girls, 54 boys 4.5 (0.8)y	**LPA**: 57.5 (37.4) min/day **MVPA**: 104.1 (60.4) min/day **TV viewing**: TV: 90.7 (50.7) min/day	Yes	No	No ≥ 60min/day Yes ≤ 120 min/day
Combined Heart Rate and Accelerometry Actiheart							
Collings et al.(2013)	England	7 day 24 h wear	398 196 girls, 202 boys Median (IQR): 4.1 (0.08) y	**LPA**: 432.6 (63.0) min/day **MVPA**: 84.7 (46.4) min/day **Sed**: 329.3 (72.7) min/day	Yes	Yes	NA
Heart Rate							
Benham-Deal (2005)	USA	Polar Vantage XL heart watch HR monitor for 12 h ≥ 3 day (including 1 weekend day) 60 s Intensity: Low: <129 bpm or <65% max HR Moderate: 130 - 159 bpm or 65 - 75% max HR High: >160bpm or >75% max HR	39 20 girls, 19 boys, 3 - 5 y 4.3 (0.7) y	No difference weekend & week Episodes lasted 5 - 10 mins: 85% of on week 76% on weekend. ≥ 60 min/day **MVPA** week: 71% weekend: 46%	Yes on week day No on weekend day	NA	NA

Measurement approach Author	**Country**	**Measurement Protocol**	**No. participants (n), gender, age range, mean (SD)**	**Key findings Unless stated figures are the mean (SD)**	**≥ 60 min/day MVPA**	**≥180 min/day TPA**	**≤60 min/day or ≤120 min/day of ST**
Jago et al. (2005)	USA	HR min by min 6 - 12 h/day 4 day at baseline 3 day at year 2 & 3	149 at baseline, followed for 3 y 3 - 4 y 149, y 1 147 y 2 138 y 3	Year 1: **Sed**: 52.9 min/h; **MVPA** 7.6 min/h; **TV viewing:** 9.7 min/h Year 2: **Sed**: 37.0 min/h **MVPA** 5.2 min/h; **TV viewing:** 10.4 min/h Year 3: **Sed**: 35.9 min/h; **MVPA** 5.8 min/day; **TV viewing:** 11.9 min/h **Sed** in year 1 predicted **TPA** in year 2 & 3.	Yes likely	NA	No likely < 60 min/day
Sallo et al. (1997)	Estonia	Whole day heart rate (HR) monitoring ≥ 4 day Mean wear time 10.5 h	54, 29 girls, 25 boys, 4 - 8 y 7.0 (0.9) y	No difference in week & weekend in **TPA** or % time HR within ranges (119 - 139 bpm, 140 - 157 bpm, >157 bpm) seen in subsample (n = 22, 12 girls, 10 boys) 86% of girls & 84% boys had 5 - 9 min of sustained MVPA (HR ≥ 139bpm) 17% girls, 20% boys had ≥ 20 min of sustained MVPA (HR ≥ 139bpm)	NA	NA	NA

AEE: Activity energy expenditure; BMR: basal metabolic rate; bpm: beats per minute; HR: heart rate; LPA: light physical activity; MVPA: moderate-to-vigorous physical activity; NA: not available; PA: physical activity; RCT: randomised controlled trial; ST: screen time; TEE: total energy expenditure; TPA: total physical activity;; VPA; vigorous physical activity; VVPA: very vigorous physical activity; cpm: counts per minute.

APPENDIX II: Ethical approval, information sheets, consent forms, CARS reliability and mechanical calibration studies

Appendix II a: Ethical approval from Queen Margaret University College ethics committee

Queen Margaret University College
EDINBURGH

Jane Hislop
Physiotherapy
School of Health Sciences

Linda Graham
Registry Officer
Queen Margaret University College
Clerwood Terrace
Edinburgh EH12 8TS

Tel: 0131 317 3219
Email: lgraham@qmuc.ac.uk

6 October 2006

Dear Jane

Ethical Approval – An investigation into free-play physical activity of preschool children in Lothian

Thank you for your email response dated 21 September 2006 to the enquiry I sent you following consideration of your application by the Research Ethics Committee.

The Committee received your response confirming your Disclosure Scotland check which you had been asked to forward and members have confirmed that they are happy to grant full ethical approval for your research.

A standard condition of this ethical approval is that you are required to notify the Committee, in advance, of any significant proposed deviation from the original protocol. Reports to the Committee are also required once the research is underway if there are any unexpected results or events that raise questions about the safety of the research. Notification of completion of the study is also required – please find the appropriate form for this enclosed.

We would like to thank you for your co-operation and wish you well with your project.

Yours sincerely,

Linda Graham

Linda Graham
Secretary to the Research Ethics Committee

Campuses at Corstorphine, Edinburgh EH12 8TS | Leith, Edinburgh EH6 8HF | Gateway Theatre, Edinburgh EH7 4AH
Telephone +44(0)131 317 3000 www.qmuc.ac.uk

THE QUEEN'S ANNIVERSARY PRIZES
2002

ESTABLISHED 1875 125 YEARS OF EXCELLENCE

Appendix II b: Permission from Edinburgh City Council to access Nurseries

·EDINBVRGH·
THE CITY OF EDINBURGH COUNCIL

CHILDREN AND FAMILIES
NEIGHBOURHOOD SERVICES

Jane Hislop
16 Wellington Street
Edinburgh
EH7 5ED

Your Ref:

Our Ref:

Date: 28 March 2006

Dear Jane,

Research Proposal - PLD

Thank you for sending your research proposal for a PLD on physical activity in pre-school children and for discussing it in detail with Carolyn Martin.

I am writing to confirm that the Children & Families Department supports the research in principle and is happy to facilitate the study by giving access to children in our nurseries at the appropriate point. We understand that what would be required would be access to a number of children from across a range of nurseries in Phase 1, due to take place from October 2006 – April 2007 and a number of children from across a range of nurseries in Phase 2, due to take place from January – June 2008.

We look forward to the outcome of the research & would agree to collaboration on the understanding that the results would be made available to us to inform the city of Edinburgh's response to what is a growing problem.

We look forward to hearing that you have the final agreement of the university to proceed.

Yours sincerely

Ian Glen
Head of Neighbourhood Services

IAN GLEN
HEAD OF NEIGHBOURHOOD SERVICES
Waterloo House, Floor2, 17 Waterloo Place EDINBURGH EH1 3BG Tel: 0131 529 2220 Fax: 0131 529 2224

Appendix II c: Information sheet for parents and carers

An investigation into free-play physical activity of preschool children in Lothian: Phase 1

Queen Margaret University

EDINBURGH

Parent/Guardian Information Sheet

PHASE 1

Title: An investigation into free-play physical activity of preschool children in Lothian

Short title: Free-play activity in preschool children

Dear Parent or Guardian,

You are being invited to consider whether you wish to give permission for your child to take part in the following study. Before you decide whether you agree to your child taking part it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully. Take time to talk to others about the study if you wish.

Please feel free to get in touch with me if you have any questions or if there is anything that is not clear and you would like more information. Take time to decide whether you wish to take part.

Yours sincerely,

Jane Hislop

Physiotherapy Lecturer
Queen Margaret University
Leith Campus
Edinburgh
EH7 5ED
jhislop@qmuc.ac.uk

Tel: 0131 317 3666

NB. In the following information sheet where the word "parent" is used, please read parent or guardian i.e. those who have parental responsibility, which may include a legal representative e.g. grandparent.

An investigation into free-play physical activity of preschool children in Lothian: Phase 1

What is the purpose of the study?
Studies suggest that young Scottish children have very low levels of physical activity. This study plans to gather information on physical activity levels in young children living within the Edinburgh and Lothian area. The study is being conducted as research for a PhD and will be in 2 stages. You are being invited to consider your child's involvement in the first stage. The first stage of this study will look at the accuracy of activity monitors (accelerometers) to measure activity levels. The second stage will use the accelerometers to collect information on children's activity levels over a 4 day period. It will then be possible to find out if children are reaching recommended levels of 60 minutes of moderate to vigorous activity per day.

Why has my child been chosen?
Preschool children who attend Edinburgh City Council Nurseries and contracted nurseries have been invited to take part. This study will involve approximately 35 children.

Does my child have to take part?
Taking part in this study is voluntary and will not affect your child's preschool provision. Your child does not have to take part and you or your child can decide that you do not wish them to participate and can withdraw at any stage without giving a reason.

What will happen to my child if we agree to take part?
Your child will be asked to wear an activity monitor (accelerometer) on a belt worn around their waist for one hour, while they take part in their usual play activities within their nursery. The monitor is a lightweight, tamper-free device which will not interfere with your child's movements. To ensure that your child is not disturbed by the presence of a researcher, the one hour play session will be video recorded. The video recording will be used to check the information on the monitor against the child's observed movements. The recording will be for the researcher's use only and will be kept in a secure manner during the study and destroyed on completion of the study. I will also measure your child's height and weight, this is to allow me to check that the accuracy of the monitor is not influence by these factors.

If you are happy to take part, and satisfied with the explanations, you will be asked to sign a consent form. If your child is able to understand the research and is happy to take part you will be asked to sign an "assent" form with them. You will be given a copy of the information sheet and the signed consent and assent forms to keep for your records.

What are the possible disadvantages and risks of taking part?
This study will present no additional risks or disadvantages to your child's normal daily activities. The study will be done within the usual nursery hours and environment.

What are the benefits of taking part?
There will be no direct benefits to you or your child from taking part in this study. The information gathered will provide useful information on how to measure physical activity levels in young children. This information is necessary so that the activity levels of young children can be measured correctly. It will then allow us to see if children are reaching the Scottish Executive's target of 60 minutes of moderate to vigorous activity per day.

J.Hislop Page 2 23/07/2007

What if there is a problem?
If something goes wrong there are no specific compensation arrangements for participating in this study. If you have any concerns about any aspect of this study, you can speak me as the researcher in the first instance. If you remain unhappy and wish to complain formally, you can do this through contacting the Registrar at Queen Margaret University or through contacting Edinburgh City Council Children and Families unit. You can also contact Gill Baer, physiotherapy lecturer at Queen Margaret University who will be acting as an independent advisor for this study (contact details are provided at the end of this form).

Disclosure Scotland
In order to undertake this study I have gone through an enhanced disclosure process with Disclosure Scotland. This is part of the Scottish Criminal Record Office and was undertaken to reassure participants and ensures enhanced security and protection for vulnerable populations in society.

Will my child's taking part in this study be kept confidential?
Yes. All information about your child's participation will be kept confidential.
All personal information which is collected about you and your child during the course of the study will be kept confidential. Video recorded information will be for the researcher and their supervisor's use only and will be kept in a secure manner during the study and destroyed on completion of the study.

What will happen to the results of the study?
The results from this study will be presented within a PhD thesis and a report for Edinburgh City Council. The results may also be published within a journal. Information will be made available to you and anonymised information will be made available to participating nurseries. You and your child will not be identified in any publication/report.

Who is organising and funding the research?
This research is organised by Jane Hislop, a Physiotherapy Lecturer and part-time PhD student at Queen Margaret University. Edinburgh City Council's Children and Families unit are supporting this study.

Who has reviewed the study?
Queen Margaret University Research Ethics Committee has reviewed this study. Edinburgh City Council has also granted approval.

Please don't hesitate to get in touch if you have any further question or wish any additional information.

Thank you for considering taking part in this study

Contact Information
Jane Hislop, Physiotherapy Lecturer
Queen Margaret University
Leith Campus
Edinburgh
EH7 5ED
Tel: 0131 317 3666
jhislop@qmuc.ac.uk

Independent Advisor:
Gill Baer
Senior Physiotherapy Lecturer
Queen Margaret University
Leith Campus
Edinburgh
EH7 5ED
Tel: 0131 317 3356
GBaer@qmuc.ac.uk

Parent or guardian will be given a copy of the information sheet to keep and a signed consent form to keep.

Information sheet dated 23rd April 2007 version 4. Study 1.

Appendix II d: Consent form for parents and carers

An investigation into free-play activity in preschool children in Lothian: Phase 1

Please return in SAE enclosed. Thank you!

Queen Margaret University

EDINBURGH

Centre Number::
Study Number:
Subject Identification Number for this study:

CONSENT FORM

Short Title: Free-play activity in preschool children

Name of Researcher: Jane Hislop

Please initial box

1. I confirm that I have read and understand the information sheet dated 23rd July 2007. version 4) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

2. I understand that my child's participation is voluntary and that they can withdraw at any time, without giving any reason, without their preschool provision or legal rights being affected. ☐

3. I understand and agree to my child being video-recorded during 1 hour of play in the nursery setting. ☐

4. I agree to my child taking part in the above study. ☐

5. I agree to the information being used for a PhD thesis and future reports/ publications ☐

Name of Parent/guardian	Date	Signature
______________________	______________	______________________
Name of Person taking consent (if different from researcher)	Date	Signature
______________________	______________	______________________
Researcher	Date	Signature
______________________	______________	______________________

When completed, 1 for parent/guardian; 1 for researcher site file

jhislop Page 1 23rd July 2007

Appendix II e: Information pamphlet for parent's to read to children

Free-play activity in preschool children

Phase 1

This information sheet can be shown/read to the child by their parent/guardian.

Hi, my name is Jane. I work for Queen Margaret University and I am writing to see if you might be able to help me with my project.

Before you say whether you would like to help I would like to tell you a bit about this project.

If you would like to talk to me before you decide you can ring me on 0131 317 3666. If I'm not there, you can leave a message on my answer phone and I'll call you back.

Jane Hislop
Physiotherapy Lecturer
Queen Margaret University
Edinburgh
Tel: 0131 317 3666
Email: jhislop@qmuc.ac.uk

Queen Margaret University
EDINBURGH

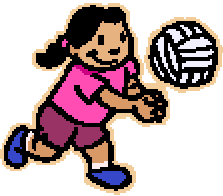
Why is this project being done?
Some people think that boys and girls don't do enough running around. I would like to find out how much running around boys and girls like you do.

If I take part what will happen?
If you say you would like to take part I would ask you to wear a belt with a small box on it for 1 hour while you are at nursery. This is a special box that can tell me how much movement you are doing. I would like to video you while you are wearing the box so I can check that the box is working correctly. I will be the only person to watch the video and this video will be destroyed at the end of the study.

Could I get hurt?
No you won't be asked to do anything different. The belt is very light and hopefully you won't even know you are wearing it!

Do I have to take part?

You don't have to take part and even if you say yes at the beginning you can change your mind and stop at anytime.



Who will know I am in the study?

Your name will not be used in the study and information about you will not have your name on it.

Who is doing the study?

The study is being done by me, Jane Hislop. I work at Queen Margaret University.

What happens next?

I will write about how much movement you and other children like you do in a report. This will let people know how much running around children like you do.

Will my name be in the research report?

No. All the names will be changed so that nobody will recognise you.

Queen Margaret University
EDINBURGH

jhislop
An investigation into free-play activity in preschool children in I

Have a think about whether you would like to take part. If you decide you want to take part your parent/guardian will be asked to sign a special form that gives your permission.

If you want to ask any more questions you or your parent/guardian can get in touch with me.

Thank you!

Jane

Jane Hislop
Physiotherapy Lecturer
Queen Margaret University
Edinburgh

Tel: 0131 317 3666
Email: jhislop@qmuc.ac.uk

Information sheet dated:
23rd April 2007 version 4

Queen Margaret University
EDINBURGH

Appendix II f: Assent form for parents and carers to read to children

An investigation into free-play activity in preschool children in Lothian

Please return in SAE enclosed. Thank you!

Queen Margaret University College

ASSENT FORM FOR CHILDREN
(to be completed by the child and their parent/guardian)

Short title: Free-play activity in young children

Child (or if unable, parent on their behalf) to circle all they agree with please:

Have you read (or had read to you) about this project?	Yes/No
Has somebody else explained this project to you?	Yes/No
Do you understand what this project is about?	Yes/No
Have you asked all the questions you want?	Yes/No
Have you had your questions answered in a way you understand?	Yes/No
Do you understand it's OK to stop taking part at any time?	Yes/No
Are you happy to take part?	Yes/No

If any answers are 'no' or you **don't** want to take part, **don't** sign your name!

If you do want to take part, please write your name and today's date

Your name ____________________

Date ____________________

Your parent or guardian must write their name here too if they are happy for you to do the project

Print Name ____________________

Sign ____________________

Date ____________________

The person who explained this project to you needs to sign too:

Print Name ____________________

Sign ____________________

Date ____________________

Thank you for your help.

jhislop Page 1 23rd July 2007

Appendix II g: CARS activities

Appendix Table II.i: Summary of description of activities and CARS codes adapted from Puhl et al. (1990)

CARS level & definition	Activities
Level 1: Stationary/Motionless (resting/motionless for 3s or more, head, finger, or foot movement only	1. Sleeping 2. Lying, standing, sitting, squatting, and kneeling 3. Floating motionless in the water 4. Riding in a wagon
Level 2: Stationary/Movement of Limb(s) or Trunk (very easy) (arm, trunk and/or leg movements without moving the entire body from one place to another	1. Standing, sitting, squatting, and kneeling with limb or trunk movement (swinging, swaying, bending, twisting, kicking, striking, throwing, digging in the sand, pantomiming song) 2. Going down a slide (requires some balance and body control) 3. Hanging or partial hanging, leaning (on fence, pole, etc.) 4. Light calisthenics, involving swinging, twisting, and stretching while seated or standing (trunk twist, arm circles, leg swings, side bending, toe touching done in an easy manner) 5. Add-on rule: Standing motionless (1) + supporting a moderately heavy object (1) = 2
Level 3: Translocation (slow/easy) (moving body from one location to another)	1. Walk/Run • walking at a leisurely or moderate pace (1 - 3 mph) • marching, skipping, hopping, jumping, crawling, rolling (slowly) • marching in one place (same as marching with translocation) 2. Rollers and Wheels • cycling • skate boarding • roller/ice skating • scooter 3. Swimming • swimming with support 4. Calisthenics • sit-ups • push-ups 5. Climbing/Sliding

CARS level & definition	Activities
	• swinging on a swing (maintaining momentum) 6. Tumbling • tumbling/wrestling (easy) 7. Add-on-rule examples: • going down a slide (Level 2) and pushing self (+1) = 3 • other stationary calisthenics (Level 2) requiring moderate effort (+1) = 3 (such as continuous up and down movements at a moderate pace)
Level 4: Translocation (medium speed/moderate) (moving from one location to another)	1. Walk/Run • walking very fast (3+ mph) • very slow jog (i.e., very slow run) • continuous skipping, hopping, jumping, leaping, crawling • walking up stairs • walking up hill 2. Rollers/Wheels • cycling • roller/ice skating • skate board • scooter 3. Swimming • swimming with minimal support • swimming very slowly without support • treading water 4. Calisthenics • moderate vigorous exercises 5. Climbing/Sliding • climbing monkey bars, fence, etc. • climbing up a slide backwards • hanging by arms from a bar with legs swinging • climbing over and under bars 6. Tumbling

CARS level & definition	Activities
	• forward rolls, backward rolls (sequence of three or more) • tumbling/wrestling, (moderate) • cartwheels, in a sequence of three or more • jumping on a trampoline 7. Add-on-rule examples • walking (Level 3) + carrying, pushing, or pulling a large or heavy object (such as pushing a merry-go-round while walking) (+1) = 4 • walking up hill or stairs (Level 3 + 1) = 4 • swinging (Level 3) + vigorous movement or gaining momentum (+1) = 4
Level 5: Translocation (fast or very fast/hard) (moving body from one location to another)	1. Walk/Run • running or jogging (fast or very fast) • climbing stairs fast or climbing using arms • walking up steep grades • fast skipping • rope jumping 2. Rollers/Wheels • cycling • skate boarding • roller/ice skating • scooter 3. Swimming • without support 4. Calisthenics • pull-up, push-up • jumping jacks • squat thrusts • continuous jumping, kicking, swinging from arms (e.g. traveling with alternating hand grasp on overhead ladder) 5. Climbing/Sliding • Swinging from arms (e.g. travelling with alternating hand grasp on overhead ladder) 6. Tumbling

CARS level & definition	Activities
	• tumbling skills in sequence • wrestling • gymnastics and apparatus 7. Add-on-rule examples • walking (Level 3) while carrying a very heavy object or a person (+2) = 5 • jogging (Level 4) + pushing a merry-go-round (+1) = 5

CARS: Children's Activity Rating Scale (Puhl et al., 1990)

Appendix II h: Reliability of CARS

Equation II i: 95% Confidence interval of mean difference

$$95\%\ CI = đ\ \pm 2\ X\ SE$$

Equation II ii: Calculation of the Standard Error (*SE*)

$$SE = (SD_{diff})/\sqrt{n}$$

Appendix Table II.ii: Mean difference in CARS and 95% confidence interval between tests.

Participant	Mean difference (đ) in CARS
1	0.09 (-0.20 to 1.14)
2	-0.21 (-0.91 to 0.48)
3	0.33 (-0.34 to 0.99)
4	-0.08 (-0.74 to 0.58)
5	0.03 (-0.40 to 0.46)
6	0.33 (-0.51 to 1.17)

Appendix Table II.iii: ICC values with 95% Confidence Intervals (CI)

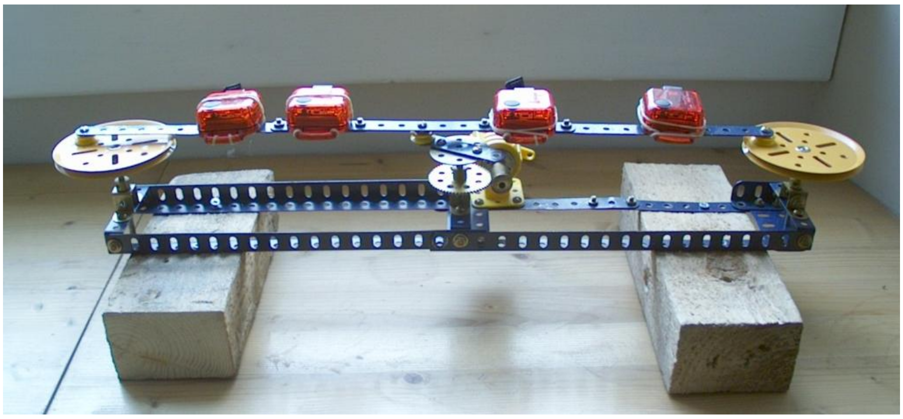
Participant	ICC	95% CI
1	0.96	0.91 to 0.98
2	0.72	0.43 to 0.86
3	0.92	0.83 to 0.96
4	0.93	0.86 to 0.95
5	0.83	0.69 to 0.90
6	0.92	0.83 to 0.96

Appendix II i: Mechanical calibration studies of accelerometers

Calibration equipment

A mechanical set up was constructed based on a calibration unit developed by Brage et al. (2003). This consisted of two larger (77 mm diameter) lateral wheels and a smaller central wheel (44mm diameter) which was driven by a motor. The wheels were connected by a horizontally placed linkage bar which could be attached in two positions which would allow for the radius positions for the lateral wheels of 25 mm (Appendix Figure II.i).

Appendix Figure II.i: Mechanical calibration unit.



The accelerometers were fixed in position on top of the connecting linkage bar along their x axis. A maximum of six GT1M/7164 or four RT3 units could be positioned on the connecting linkage bar allowing the accelerometers to be tested simultaneously in batches of either four or six. Rotation of the wheels caused the connecting linkage bar to move the accelerometers through a sinusoidal motion along two axes (anteroposterior and mediolateral) providing acceleration on each of these axis. The central wheel was driven by a motor fed by a power supply (Farnell UK Limited, Leeds, UK). Changes in the frequency of movement of the wheels, and thus the connecting linkage bar, was achieved by altering the voltage output from the power supply. An incremental increase in 0.25V from the power supply resulted in an average incremental increase in frequency of 0.2 Hz. The range of frequencies ranged from 0.5 Hz to 3.3 Hz. This meant that by using the different frequencies the accelerometers could be subjected to a total of 14 different acceleration settings.

The average acceleration (ā) produced by the mechanical set-up during testing was calculated using the equation (Equation II: iii) outlined in the study by Brage et al. (2003), where acceleration is a function of radius (r) and frequency (f) of oscillation.

Equation II iii: Calculation of average acceleration sinusoidal oscillator.

$$\bar{a} = 8 \cdot \pi \cdot r \cdot f^2$$

(Brage et al. 2003)

Appendix II j: Calibration test 1: Position test

Aim

The aim of this study was to determine if there was uniformity of motion between the different locations on the mechanical set up.

Method

To determine if there was uniformity of motion between the different locations on the connecting linkage bar the accelerometers were tested in each of the six test positions. Initially, ten GT1M accelerometers were available for testing however one accelerometer was excluded prior to data collection as it was found to be faulty and not charging. Nine GT1M accelerometers were tested three times in each of the six possible positions along the horizontal linkage at a low (1 Hz) and high frequency (3Hz) using the radius setting of 25 mm. This corresponded to peak acceleration at 1.06 $m \cdot s^{-2}$ and 9.59 $m \cdot s^{-2}$ and average acceleration of 0.63 $m \cdot s^{-2}$ and 5.65 $m \cdot s^{-2}$.

The mechanical unit was set to run for one minute pre-test and testing took place for each frequency for three minutes. The middle 2 minutes for each set of data collected was used for analysis, so that the first and last 30 seconds were excluded from analysis. A one minute interval was taken between test change in frequency and position. The accelerometers were set to record data in 1-s epochs.

A repeated measures ANOVA was carried out to test whether there was a significant difference between the output of the accelerometers in the different positions on the connecting linkage bar. A repeated measures ANOVA depends on the assumption of sphericity, which refers to the equality of variance of the difference between 'treatment' conditions (Field 2012), which in the current study refers to the variance of difference in accelerometry output in counts per second between the six positions on the mechanical set-up.

The mean score (SD) for each monitor was calculated at 1 Hz and 3 Hz. The ICC and 95% CI was calculated to determine whether the activity scores were consisted between tests.

Results

The mean (SD) of the counts/second of the accelerometers are presented in Appendix Table II.iv.

Appendix Table II.iv: Mean (SD) of counts/second for GT1M accelerometers during mechanical calibration.

Frequency	Position 1	Position 2	Position 3	Position 4	Position 5	Position 6
3 Hz	4830.52 (87.34)	4848.73 (93.61)	4889.95 (131.29)	4818.88 (225.91)	4818.09 (242.18)	4813.40 (219.15)
1 Hz	1030.07 (432.23)	1021.97 (398.14)	840.14 (222.23)	929.07 (425.22)	1037 (375.15)	1008.88 (438.85)

At 3 Hz the Mauchley's test indicated that the assumption of sphericity had been violated (χ^2 (14) = 26.20, p < 0.05); therefore the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (ϵ = 0.42) giving a more conservative measure of the *F*-ratio. The results suggested that the counts per second were not significantly affected by the position of the accelerometer on the horizontal linkage bar, F (2.10,12.58) = 0.57, p > 0.59, ω^2 = 0.01. The *F*-ratio (*F*), which is a measure of the ratio of variation explained by the model with the variation explained by systematic factors, while the effect size (ω^2) is the magnitude of the observed effect (Field 2012).

At 1 Hz the Mauchley's test suggested that sphericity had been not been violated (χ^2(14) = 5.71, p > 0.98). With sphericity assumed therefore the results suggest that counts per second were not significantly affected by the position of the accelerometer on the horizontal linkage bar, F (5,30) = 0.26, p > 0.93 ω^2=0.02.

The average ICC value was 0.93 which together with the 95% CI (0.79 to 0.99) being close to 1 suggested 'excellent' agreement between the units.

Discussion and conclusion

At both 3 Hz and 1 Hz, the repeated measures ANOVA revealed no significant location effect between the different positions on the mechanical set up. The *F*-value value was < 1, which suggests that the unsystematic ratio was greater, and therefore there was more unsystematic variation than systematic variation, i.e. the experimental effect, in this case the effect of position. This allows for the accelerometers to be evaluated in the subsequent experiments without the results being influenced by systematic experimental effect from the mechanical set up.

Appendix II k: Calibration test 2: Intra-unit and inter model reliability of RT3, GT1M, 7164

Aim

The aim of this study was to assess the intra-unit reliability of each of the individual accelerometers and to assesses the inter model reliability of the accelerometer of the same model. This was undertaken pre-field data collection and post-field data collection for studies in Chapters four to seven.

Method

To test for intra-unit agreement nine GT1M, eight 7164 accelerometers and ten RT3 accelerometers were tested. The GT1M and 7164 were tested in batches of six, while the slightly larger and heavier RT3 was tested in batches of four. Batches of accelerometers of the same model were tested three times at three different frequencies using 25 mm radius setting: 1 Hz, 2 Hz and 3 Hz. The accelerometers were therefore subjected to 1.5, 4.3 and 9.6 $m{\cdot}s^{-2}$ of peak acceleration respectively and 0.63, 2.51 and 5.65 $m{\cdot}s^{-2}$ of average acceleration. Each test lasted 5 minutes and the GT1M and 7164 accelerometers were set to record acceleration in 1-s epochs. The RT3 was set to record at 1-s epochs on Vector Magnitude mode (VM) which combines counts from three axes.

Similar to other studies (Rothney et al. 2008) the accelerometers were positioned along their sensitive axis, so that the GT1M and 7164 were position along their x-axis and the RT3 positioned so that they were oscillated in an anteroposterior direction along their x-axis and a mediolateral direction along their y-axis (Esliger and Tremblay 2006).

The data were transferred into an Excel spreadsheet then exported to SPSS (version 15.0) for analysis. The first and last minutes of data were excluded so that the middle 3 minutes of data were used. To determine intra-unit variability the standard error of measurement (*SEM*) and coefficient of variation (CV) were calculated for each of the frequencies. To calculate the CV, the standard deviations of the means were divided by the corresponding unit means (Equation II iii).

Equation II iii: Calculation of coefficient of variation.

$$\mathrm{CV} = 100\left(\frac{SD}{\bar{\mathrm{x}}}\right)$$

(Munro 2001)

To determine the inter-model variability between like-model accelerometers (e.g. between the GT1M units, the 7164 units and between the RT3 units), the CV% was calculated and plotted to allow comparison between the inter-unit reliability of the different models. An ICC with a two-way random effects model (3,1) together with 95% CI was also calculated using SPSS as recommended Batterham et al. (2000).

Results

During the field data collection, one of the RT3 accelerometers stopped charging and had to be excluded from the research. Ten RT3s were therefore tested pre-field data collection and nine were tested post-field data collection. Two 7164 failed during calibration testing (data output was 0) and both were excluded from the analysis. Six 7164 were included in the pre-and post-field data collection tests. Nine GT1M accelerometers were tested in the pre- and post-field data collection.

A summary of the results of the intra-unit variability for the GT1M, 7164, and RT3 accelerometers, pre- and post-field data collection are presented in Appendix Table II.v, Appendix Table II.vi and Appendix Table II.vii respectively. Appendix Table II.viii presents the mean (CV)% for the different models of accelerometer.

Appendix Table II.v: Intra-unit reliability GT1M.

	Pre: SEM	**Pre: CV%**	**Post: *SEM***	**Post: CV%**
GT1M: 1	0.23	0.73	0.38	1.72
GT1M: 2	0.06	0.82	0.19	1.62
GT1M: 3	0.16	0.72	0.27	1.49
GT1M: 4	0.05	0.56	0.31	1.34
GT1M: 5	0.16	0.59	0.17	1.11
GT1M: 6	0.26	0.93	0.27	1.90
GT1M: 7	0.08	0.45	0.31	1.44
GT1M: 8	0.08	0.60	0.29	1.56
GT1M: 9	0.07	0.56	0.66	1.97

Appendix Table II.vi: Intra-unit reliability 7164.

	Pre: SEM	Pre: CV%	Post: *SEM*	Post: CV%
7164: 1	1.37	6.41	0.64	4.65
7164: 2	1.97	7.70	0.59	4.54
7164: 3	2.12	8.76	0.54	4.58
7164: 4	2.61	8.64	0.71	5.04
7164: 5	3.11	9.61	0.53	4.63
7164: 6	2.43	8.58	0.55	5.07

Appendix Table II.vii: Intra-unit reliability RT3.

	Pre: SEM	Pre: CV%	Post: *SEM*	Post: CV%
RT3: 1	1.42	5.66	1.61	8.05
RT3: 2	1.94	5.90	1.30	2.61
RT3: 3	1.50	4.33	1.67	7.21
RT3: 4	1.73	4.78	N/A	N/A
RT3: 5	0.95	3.91	1.50	4.40
RT3: 6	0.92	5.70	1.30	5.44
RT3: 7	1.07	5.17	1.70	5.09
RT3: 8	1.64	5.81	0.70	5.52
RT3: 9	0.93	4.95	1.44	5.42
RT3: 10	1.11	5.70	1.20	5.60

Appendix Table II.viii: The intra-unit reliability of individual accelerometers by model pre- and post-field data collection.

	Mean CV (%) with range	
Accelerometry model	**Pre**	**Post**
RT3	5.19 (3.95 - 5.99)	5.93 (2.89 - 9.58)
GT1M	0.66 (0.45 - 0.82)	1.57 (0.36 - 3.38)
7164	8.28 (8.28 - 9.6)	4.75 (4.54 - 5.06)

CV coefficient of variation.

Appendix Table II.ix presents a summary of the mean the inter-unit reliability of the different accelerometers of the same model, comparing pre- and post-field data collection.

Appendix Table II.ix: The inter-unit reliability of different accelerometers of the same models pre- and post-field data collection

	Mean CV (%) with range	
Accelerometry model	**Pre**	**Post**
RT3	7.79 (4.24 - 11.97)	14.14 (4.84 - 23.56)
GT1M	3.29 (1.91 - 4.89)	2.39 (1.94 - 2.89)
7164	9.2 (8.63 - 10.85)	6.53 (5.75 - 7.97)

Appendix Figure II.ii presents the plots of the CV% for each of the accelerometer models at each of the frequencies (converted to average acceleration) to allow visual cross comparison of the inter-unit of the different models during pre- and post-field data collection.

Appendix Figure II.ii: Inter-monitor coefficient of variability (CV)% for the RT3, GT1M and 7164 pre-data collection.

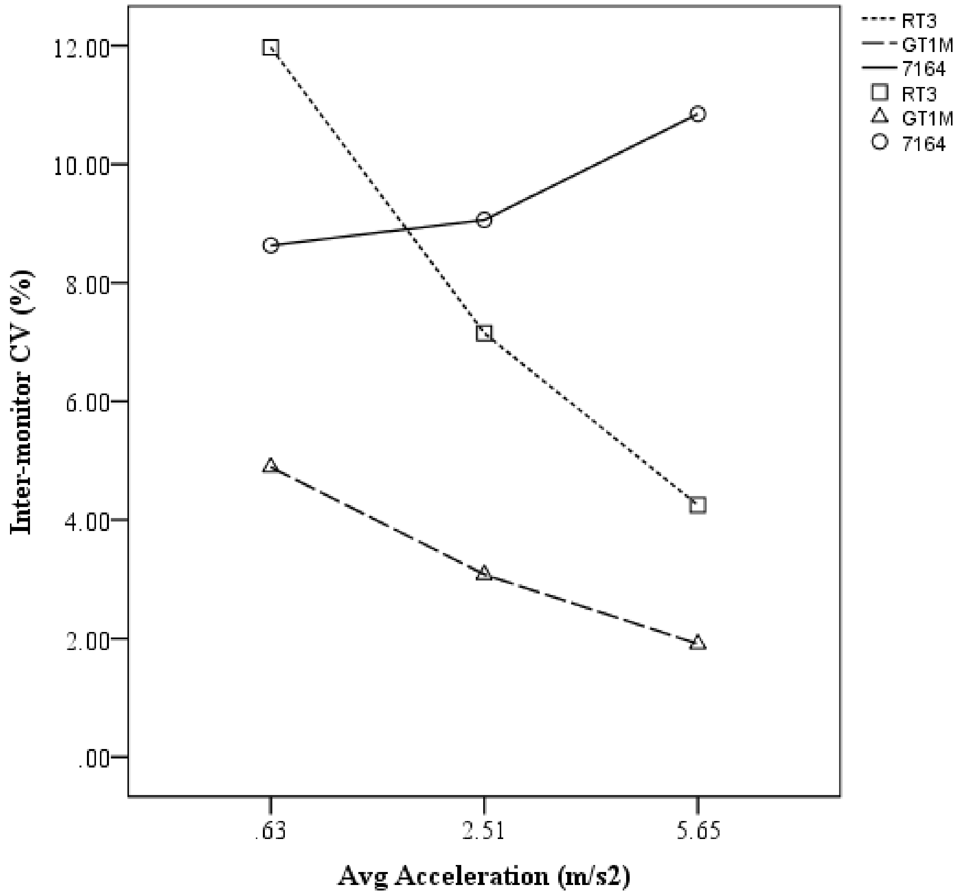
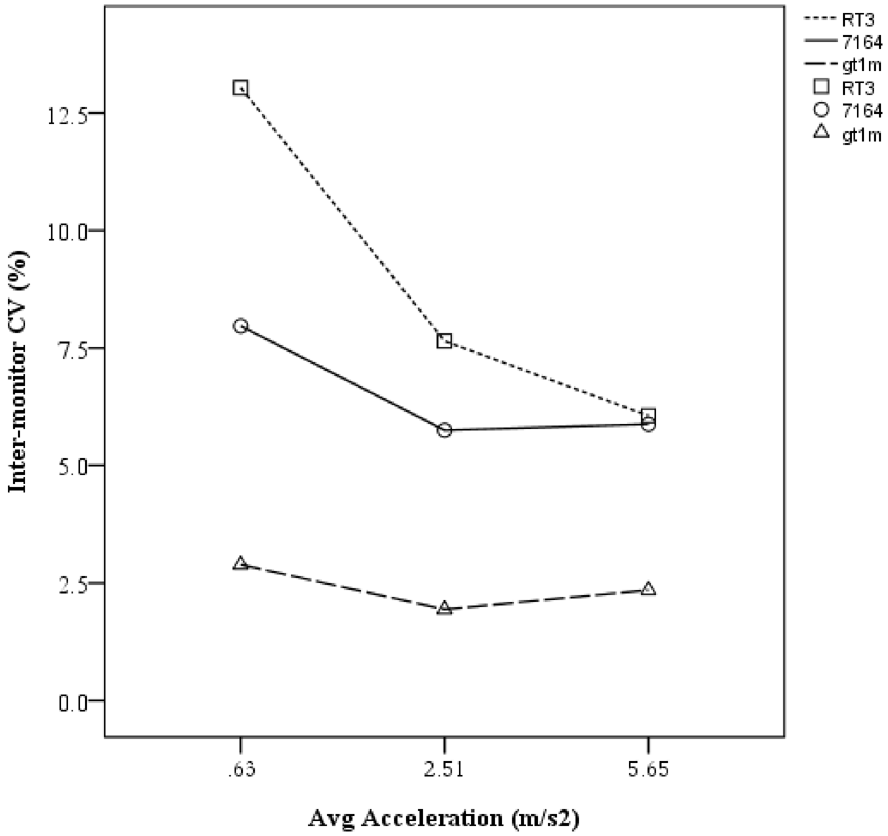


Figure II:c: Inter-monitor coefficient of variability (CV)% for the RT3, GT1M and 7164 post field data collection.



Appendix Table II.x presents the results of the ICC for inter-model reliability between accelerometers of the same model.

Appendix Table II.x: Interclass correlation coefficient (ICC) between units of the same model pre- and post-field data collection.

	ICC (95% Confidence Interval)	
Accelerometer model	**Pre**	**Post**
RT3	0.98 (0.97 - 0.99)	1.00 (0.99 - 1.00)
GT1M	0.99 (0.97 - 0.99)	0.99 (0.98 - 0.99)
7164	0.99 (0.95 - 1.00)	1.00 (0.99 - 1.00)

Discussion

The results suggested that the GT1M models had good intra-unit reliability of individual accelerometers as well as good inter-unit reliability between accelerometers of the same model with ICCs that were very close to 1 (perfect agreement). The RT3 and 7164 both had higher CV values, particularly at lower accelerations in comparison to the GT1M accelerometer. These findings are similar to the values reported in other mechanical calibration studies (Rothney et al. 2008). The CV% values and the ICC values were similar both pre-and post-data collection.

Appendix II l: Calibration test 3: Comparison between 7164, GT1M and RT3

Aim

The aim of this final calibration study was to compare different models of accelerometers during mechanical calibration.

Method

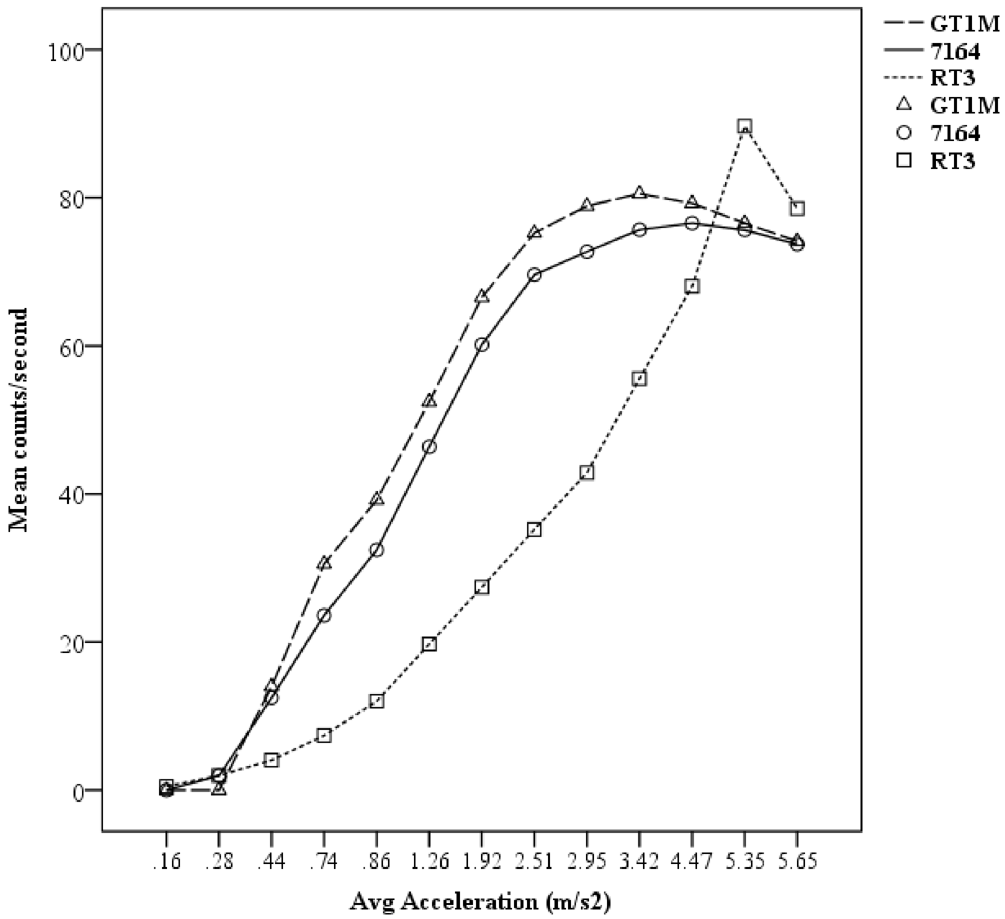
To compare accelerometer models the RT3, GT1M and 7164 were simultaneously oscillated on the calibration unit subjecting them to 14 accelerations between 0.5 to 3.3 Hz in average increments of 0.2 Hz. This corresponds to average acceleration ranging from 0.6 to 5.65 $m \cdot s^2$. Each test lasted 5 minutes and accelerometers were set to record acceleration in one second epochs. The middle 3 minutes of data was used for analysis. The mean of the two accelerometers of each model was plotted. Accelerometers were set to collect data in 1-s epochs.

The mean counts per second for each frequency were calculated and plotted to allow visual comparison between the accelerometer models.

Results

Appendix Figure II.iii: presents the mean counts per second at each acceleration.

Appendix Figure II.iii: comparison of counts per second between accelerometry models.



Discussion

The 7164 and the GT1M accelerometer had similar responses to increasing acceleration and the non-linear response of the accelerometers at higher accelerations has been reported in earlier mechanical calibration studies (Brage et al. 2003). The levelling off of the output at higher accelerations is thought to be an artefact of the internal filter within the accelerometers which the manufacturers have developed to detect human movement and exclude motion from other sources (Manufacturing Technology Inc. 2001).

Appendix II m: Synchronisation of GT1M accelerometers

Example of raw count data from mechanical calibration illustrating the start and finish times of the mechanical calibration set-up against the data collected by GT1M Appendix Table II.xi) and 7164 (Appendix Table II.xii) accelerometers

Highlighted in yellow is the time the mechanical unit was started or stopped this was synchronised with the PC. Highlighted in grey are the start and finish times for the GT1M/7164 accelerometers which were also synchronised with the same PC. There were six GT1M units (GT1M-1 to GT1M-6) and six 7164 units (7164-1 to 7164-6).

Appendix Table II.xi: Mechanical calibration of the GT1M.

Time	**GT1M-1**	**GT1M-2**	**GT1M-3**	**GT1M-4**	**GT1M-5**	**GT1M-6**	Calibration unit started/stopped: time synced with PC
10:30:01	0	0	60	32	46	34	**START**
10:30:02	60	35	49	54	54	49	
10:30:03	44	49	48	48	50	50	
10:30:04	43	57	50	46	49	51	
10:30:05	55	53	55	47	55	45	
10:30:06	52	48	51	54	45	51	
10:30:07	51	51	47	51	47	52	
10:30:08	52	58	52	54	47	51	
10:30:09	49	49	50	51	54	52	
10:30:10	55	45	49	54	46	46	
10:30:11	46	54	50	50	42	57	
10:30:12	47	54	55	45	46	51	
10:30:13	53	47	48	47	51	44	
10:30:14	56	49	48	52	48	45	
10:30:15	45	51	44	49	52	53	
Data continues to be collected in 1-s epochs							
10:32:30	50	48	50	47	49	56	
10:32:31	48	53	54	49	48	49	
10:32:32	51	58	57	52	55	43	
10:32:33	54	51	53	55	47	44	
10:32:34	54	48	44	48	45	52	

10:32:35	53	53	50	44	46	42	
10:32:36	46	54	55	47	50	43	
10:32:37	52	57	60	49	51	45	
10:32:38	54	50	49	49	42	54	
10:32:39	57	48	49	47	45	53	
10:32:40	49	51	55	45	52	48	
10:32:41	50	50	52	53	54	48	
10:32:42	53	4	16	50	50	55	
10:32:43	48	0	0	51	45	59	
10:32:44	0	0	0	47	21	44	
10:32:45	0	0	0	42	0	1	
10:32:46	0	0	0	1	0	0	
10:32:47	0	0	0	0	0	0	
10:32:48	0	0	0	0	0	0	
10:32:49	0	0	0	0	0	0	
10:32:50	0	0	0	0	0	0	
10:32:51	0	0	0	0	0	0	
10:32:52	0	0	0	0	0	0	
10:32:53	0	0	0	0	0	0	
10:32:54	0	0	0	0	0	0	
10:32:55	0	0	0	0	0	0	
10:32:56	0	0	0	0	0	0	
10:32:57	0	0	0	0	0	0	
10:32:58	0	0	0	0	0	0	
10:32:59	0	0	0	0	0	0	
10:33:00	0	0	0	0	0	0	**STOP**
10:33:01	0	0	0	0	0	0	
10:33:02	0	0	0	0	0	0	
10:33:03	0	0	0	0	0	0	
10:33:04	0	0	0	0	0	0	
10:33:05	0	0	0	0	0	0	
10:33:06	0	0	0	0	0	0	
10:33:07	0	0	0	0	0	0	
10:33:08	0	0	0	0	0	0	
10:33:09	0	0	0	0	0	0	
10:33:10	0	0	0	0	0	0	

10:33:11	0	12	0	0	0	0	
10:33:12	0	51	38	0	0	0	
10:33:13	7	37	46	0	0	0	
10:33:14	45	14	5	0	28	4	
10:33:15	39	10	36	10	48	51	
10:33:16	12	15	13	48	27	39	
10:33:17	24	14	16	27	21	21	
10:33:18	17	22	14	19	19	33	
10:33:19	14	15	18	28	17	18	
10:33:20	17	17	15	20	12	13	
10:33:21	15	17	15	15	24	25	
10:33:22	16	28	15	19	12	19	
10:33:23	21	19	17	19	17	20	
10:33:24	20	14	23	16	15	15	
10:33:25	18	12	24	16	15	19	
10:33:26	16	14	24	18	15	16	
10:33:27	16	16	19	22	12	19	
10:33:28	19	22	13	22	13	18	
10:33:29	19	21	24	13	13	18	
10:33:30	17	18	29	22	19	20	**START**
10:33:31	19	18	16	16	16	14	
10:33:32	14	13	19	16	18	19	
10:33:33	14	20	13	19	17	17	
10:33:34	13	20	11	18	17	16	
10:33:35	22	14	14	19	15	21	
10:33:36	22	10	21	20	16	19	
10:33:37	20	13	22	13	22	25	
10:33:38	19	15	23	18	12	18	
10:33:39	17	18	17	16	13	15	
10:33:40	17	17	17	21	15	17	
10:33:41	17	21	15	18	7	19	
10:33:42	17	19	21	16	16	15	
10:33:43	20	16	19	10	17	21	
10:33:44	21	16	18	14	15	16	
10:33:45	17	21	16	14	11	16	
Data continues to be collected in 1-s epochs							

10:36:00	16	16	17	20	12	16	
10:36:01	16	16	18	20	19	19	
10:36:02	16	11	22	20	17	17	
10:36:03	21	22	21	16	20	20	
10:36:04	16	18	25	16	20	18	
10:36:05	12	20	16	21	17	19	
10:36:06	13	18	17	20	12	13	
10:36:07	21	15	11	17	18	15	
10:36:08	18	20	15	17	13	21	
10:36:09	25	19	11	17	19	20	
10:36:10	23	13	15	17	15	19	
10:36:11	16	14	19	18	18	11	
10:36:12	10	0	0	19	16	13	
10:36:13	16	0	0	22	16	18	
10:36:14	0	0	0	27	4	17	
10:36:15	0	0	0	9	0	0	
10:36:16	0	0	0	0	0	0	
10:36:17	0	0	0	0	0	0	
10:36:18	0	0	0	0	0	0	
10:36:19	0	0	0	0	0	0	
10:36:20	0	0	0	0	0	0	
10:36:21	0	0	0	0	0	0	
10:36:22	0	0	0	0	0	0	
10:36:23	0	0	0	0	0	0	
10:36:24	0	0	0	0	0	0	
10:36:25	0	0	0	0	0	0	
10:36:26	0	0	0	0	0	0	
10:36:27	0	0	0	0	0	0	
10:36:28	0	0	0	0	0	0	
10:36:29	0	0	0	0	0	0	
10:36:30	0	0	0	0	0	0	**STOP**
10:36:31	0	0	0	0	0	0	
10:36:32	0	0	0	0	0	0	
10:36:33	0	0	0	0	0	0	
10:36:34	0	0	0	0	0	0	
10:36:35	0	0	0	0	0	0	

10:36:36	0	0	0	0	0	0	
10:36:37	0	0	0	0	0	0	
10:36:38	0	0	0	0	0	0	
10:36:39	0	0	0	0	0	0	
10:36:40	0	0	0	0	0	0	
10:36:41	0	0	0	0	0	0	
10:36:42	0	17	3	0	0	0	
10:36:43	0	21	17	0	0	0	
10:36:44	9	19	17	0	2	0	
10:36:45	18	13	16	0	18	11	
10:36:46	19	8	24	7	14	24	
10:36:47	19	19	12	15	14	17	
10:36:48	17	24	13	20	9	13	
10:36:49	20	19	15	14	1	19	
10:36:50	24	13	16	12	17	28	
10:36:51	26	12	19	9	18	21	
10:36:52	15	16	17	16	17	17	
10:36:53	10	18	13	20	13	16	
10:36:54	18	17	20	20	15	16	
10:36:55	19	14	22	17	17	17	
10:36:56	24	12	21	10	15	14	
10:36:57	21	10	28	11	14	17	
10:36:58	12	15	19	13	16	18	
10:36:59	14	17	13	15	12	18	
10:37:00	17	18	18	20	12	16	**START**
10:37:01	21	15	14	19	17	25	
10:37:02	21	14	24	15	18	22	
10:37:03	12	12	21	14	17	21	
10:37:04	11	16	17	17	11	16	
10:37:05	16	14	12	21	6	14	
10:37:06	17	22	18	20	13	14	
10:37:07	21	14	18	18	15	16	
10:37:08	15	15	22	12	19	15	
10:37:09	16	16	16	15	15	13	
10:37:10	16	14	15	20	16	15	
10:37:11	18	19	16	20	16	13	

10:37:12	17	19	18	18	16	19	
10:37:13	23	19	13	14	18	17	

Appendix Table II.xii: Mechanical calibration of the 7164

	7164-1	7164-2	7164-3	7164-4	7164-5	7164-6	
12:00:01	81	82	79	94	99	89	**START**
12:00:02	72	81	78	94	101	88	
12:00:03	74	84	76	90	92	80	
12:00:04	80	76	71	86	92	85	
12:00:05	77	77	79	95	100	90	
12:00:06	76	85	77	95	99	88	
12:00:07	70	82	78	94	100	88	
12:00:08	76	81	74	87	91	81	
12:00:09	78	75	72	86	94	84	
12:00:10	79	77	76	92	101	90	
12:00:11	79	83	75	94	98	88	
12:00:12	71	80	78	95	100	90	
12:00:13	75	83	76	91	91	81	
12:00:14	78	75	70	86	92	84	
12:00:15	76	75	77	91	102	89	
Data continues to be collected in 1-s epochs							
12:04:45	75	80	71	87	96	84	
12:04:46	72	83	73	92	92	81	
12:04:47	72	82	73	94	99	88	
12:04:48	78	75	81	92	104	90	
12:04:49	79	77	79	88	95	83	
12:04:50	75	79	71	85	95	85	
12:04:51	73	84	72	93	94	81	
12:04:52	71	79	73	98	101	88	
12:04:53	77	73	79	88	102	91	
12:04:54	81	76	77	87	96	84	
12:04:55	75	77	72	88	95	84	
12:04:56	72	84	73	92	93	82	
12:04:57	73	80	73	94	99	88	
12:04:58	78	74	81	93	104	91	
12:04:59	77	77	77	89	95	83	
12:05:00	28	21	61	21	42	39	**STOP**
12:05:01	0	0	8	0	0	0	
12:05:02	0	0	0	0	0	0	

12:05:03	0	0	0	0	0	0	
12:05:04	0	0	0	0	0	0	
12:05:05	0	0	0	0	0	0	
12:05:06	0	0	0	0	0	0	
12:05:07	0	0	0	0	0	0	
12:05:08	0	0	0	0	0	0	
12:05:09	0	0	0	0	0	0	
12:05:10	0	0	0	0	0	0	
12:05:11	0	0	0	0	0	0	
12:05:12	0	0	0	0	0	0	
12:05:13	0	0	0	0	0	0	
12:05:14	0	0	0	0	0	0	
12:05:15	0	0	0	0	0	0	
Data continues to be collected in 1-s epochs							
12:05:45	0	0	0	0	0	0	
12:05:46	0	0	0	0	0	0	
12:05:47	0	0	0	0	0	0	
12:05:48	0	0	0	0	0	0	
12:05:49	0	0	0	0	0	0	
12:05:50	0	0	0	0	0	0	
12:05:51	0	0	0	0	0	0	
12:05:52	0	0	0	0	0	0	
12:05:53	0	0	0	0	0	0	
12:05:54	0	0	0	0	0	0	
12:05:55	0	0	0	0	0	0	
12:05:56	0	0	0	0	0	0	
12:05:57	0	0	0	0	0	0	
12:05:58	0	0	0	0	0	0	
12:05:59	14	15	0	16	7	7	
12:06:00	81	88	40	97	96	87	**START**
12:06:01	75	81	82	90	98	86	
12:06:02	74	83	73	92	95	82	
12:06:03	74	82	74	94	98	88	
12:06:04	74	77	78	93	101	88	
12:06:05	73	78	76	92	100	88	
12:06:06	77	76	79	88	100	88	

12:06:07	79	76	77	87	97	86	
12:06:08	78	77	75	89	95	85	
12:06:09	77	77	74	93	95	83	
12:06:10	75	78	73	90	96	85	
12:06:11	73	79	75	89	97	87	
12:06:12	73	80	74	89	96	83	
12:06:13	74	80	75	94	96	84	
12:06:14	74	81	74	91	99	89	
12:06:15	72	79	77	93	100	89	
Data continues to be collected in 1-s epochs							
12:10:45	80	81	72	88	92	84	
12:10:46	73	83	74	88	94	82	
12:10:47	72	74	73	96	103	91	
12:10:48	76	76	81	88	104	89	
12:10:49	80	80	73	87	94	82	
12:10:50	70	84	72	91	93	81	
12:10:51	72	76	75	96	102	93	
12:10:52	79	74	78	87	104	89	
12:10:53	80	81	73	85	93	82	
12:10:54	74	84	71	91	91	83	
12:10:55	70	71	78	96	104	93	
12:10:56	80	76	80	83	96	83	
12:10:57	76	84	71	87	95	84	
12:10:58	71	82	72	94	93	84	
12:10:59	76	71	79	82	101	90	
12:11:00	19	14	59	19	35	30	**STOP**
12:11:01	0	0	4	0	0	0	
12:11:02	0	0	0	0	0	0	
12:11:03	0	0	0	0	0	0	

APPENDIX III: Results of the normality tests Chapter 3

Appendix Table III.i: Results of the normality tests: Chapter 3.

Variable	***df***	**Shapiro-Wilks test (n<50) statistic**	***p*-value**
Sed 15-s epoch	32	0.96	0.00
Sed 30-s epoch	32	0.97	0.00
Sed 60-s epoch	32	0.97	0.00
MPA 15-s epoch	32	0.90	0.00
MPA 30-s epoch	32	0.88	0.00
MPA 60-s epoch	32	0.84	0.00
VPA 15-s epoch	32	0.73	0.00
VPA 30-s epoch	32	0.60	0.00
VPA 60-s epoch	32	0.46	0.00
MVPA 15-s epoch	32	0.91	0.00
MVPA 30-s epoch	32	0.88	0.00
MVPA 60-s epoch	32	0.85	0.00

df: degrees of freedom; MPA: moderate physical activity; MVPA: moderate-to-vigorous physical activity; Sed: Sedentary behaviour; TPA: total physical activity; VPA: vigorous physical activity.

APPENDIX IV: Results of the normality tests Chapter 4

Appendix Table IV.i: Results of the normality tests Chapter 4.

Variable	***df***	**Shapiro-Wilks test (n<50) statistic**	***p*-value**
Sed 1-s epoch	31	0.85	0.001
Sed 5-s epoch	31	0.83	0.001
Sed 15-s epoch	31	0.83	0.001
Sed 30-s epoch	31	0.81	0.001
Sed 60-s epoch	31	0.81	0.001
LPA 1-s epoch	31	0.89	0.004
LPA 5-s epoch	31	0.95	0.134
LPA 15-s epoch	31	0.98	0.736
LPA 30-s epoch	31	0.97	0.621
LPA 60-s epoch	31	0.96	0.212
VPA 1-s epoch	31	0.93	0.041
VPA 5-s epoch	31	0.82	0.001
VPA 15-s epoch	31	0.62	0.001
VPA 30-s epoch	31	0.53	0.001
VPA 60-s epoch	31	0.47	0.001
MVPA 1-s epoch	31	0.73	0.001
MVPA 5-s epoch	31	0.71	0.001
MVPA 15-s epoch	31	0.71	0.001
MVPA 30-s epoch	31	0.71	0.001
MVPA 60-s epoch	31	0.62	0.001
TPA 1-s epoch	31	0.82	0.001
TPA 5-s epoch	31	0.86	0.001
TPA 15-s epoch	31	0.89	0.004
TPA 30-s epoch	31	0.88	0.002
TPA 60-s epoch	31	0.91	0.012

df: degrees of freedom; LPA: light physical activity; MVPA: moderate-to-vigorous physical activity; Sed: Sedentary behaviour; TPA: total physical activity VPA: vigorous physical activity.

APPENDIX V: Results of the normality tests Chapter 5

Appendix Table V.i: Results of the normality tests Chapter 5.

Variable	***df***	**Shapiro-Wilks (n<50) test**	***p*-value**
RT3 cpm	31	0.95	0.146
GT1M cpm	31	0.98	0.687
RT3 MVPA^{vh}	31	0.87	0.002
RT3 MVPA^{ro}	31	0.87	0.001
RT3^{WR} MVPA	31	0.82	0.001
RT3 ^{LJ} MVPA	31	0.73	0.00
RT3 MVPA	31	0.82	0.00
GT1M MVPA^{ev}	31	0.88	0.002
GT1M MVPA^{f}	31	0.91	0.012
GT1M MVPA^{pa}	31	0.89	0.005
GT1M MVPA^{pu}	31	0.91	0.012
GT1M MVPA^{s}	31	0.71	0.001
GT1M MVPA^{va}	31	0.87	0.001

Ch: Chu et al. (2007)*; df: degrees of freedom; ev: Evenson et al.* (2008)*; f: Freedson et al.* (1997)*; MVPA: moderate-to-vigorous physical activity; pa: Pate et al.* (2006)*; pu: Puyau et al.* (2002)*; ro: Rowlands et al.* (2004)*; s: Sirard et al.* (2005)*; RT3 ^{LJ}: light jog; RT3^{WR} walking relaxed; vh: Vanhelst et al.* (2010a)*; va: Van Cauwenberghe et al.* (2011)

APPENDIX VI: Results of the normality tests Chapter 6

Appendix Table VI.i: Results of the normality tests Chapter 6.

Variable	***df***	**Shapiro-Wilks (n<50) test**	***p*-value**
Total mins	31	0.891	0.004
RT3 cpm	31	0.95	0.15
GT1M cpm	31	0.98	0.69
Sed^{va}	31	0.82	0.00
Sed^{pu}	31	0.76	0.00
Seds	31	0.83	0.00
Sed^{r}	31	0.79	0.00
Sed^{ev}	31	0.72	0.00
LPA^{va}	31	0.96	0.23
LPA^{pu}	31	0.93	0.03
LPA^{s}	31	0.98	0.74
MVPA^{va}	31	0.87	0.001
MVPA^{pa}	31	0.89	0.005
MVPA^{pu}	31	0.91	0.012
MVPA^{s}	31	0.71	0.00
TPA^{pu}	31	0.91	0.011
TPA^{s}	31	0.89	0.004
TPA^{va}	31	0.92	0.008
TPA CARS	31	0.90	0.006

df: degrees of freedom; ev: Evenson et al. (2008)*; MVPA: moderate-to-vigorous physical activity; pa: Pate et al.* (2006)*; pu: Puyau et al.* (2002)*; r: Reilly et al.* (2003)*; s: Sirard et al.* (2005)*; va: Van Cauwenberghe et al.* (2011).

Appendix Table VI.ii: Results of the Wilcoxon signed rank test. *z*-scores and effect size (r)

	z-score	*p-value*	ES (*r*)
Sed^{pu} versus CARS	-0.4	0.7	-0.05
Sed^{s} versus CARS	-4.9	0.00*	-0.6
Sed^{r} versus CARS	-2.5	0.01	-0.3
Sed^{ev} versus CARS	-4.8	0.00*	-0.6
Sed^{va} versus CARS	-4.5	0.00*	-0.6
$Light^{pu}$ versus CARS	-2.3	0.02	-0.3
$Light^{s}$ versus CARS	-4.5	0.02	-0.6
$Light^{va}$ versus	-4.9	0.00*	-0.6
$MVPA^{pu}$ versus CARS	-1.9	0.05	-0.2
$MVPA^{s}$ versus CARS	-1.3	0.2	-0.2
$MVPA^{va}$ versus CARS	-3.6	0.00*	-0.5
$MVPA^{pa}$ versus CARS	-4.8	0.00*	-0.6

*ES: effect size; p: level of significance; * p <0.01; CARS: Children's Activity Rating Scale, MVPA: moderate-to-vigorous activity. s: Sirard et al.* (2005)*; pu: Puyau et al.* (2002)*; va: Van Cauwenberghe et al.* (2011)*, r: Reilly et al.* (2003)*, ev: Evenson et al.* (2008)*, pa: Pate et al.* (2006).

APPENDIX VII: Results of the normality tests Chapter 7

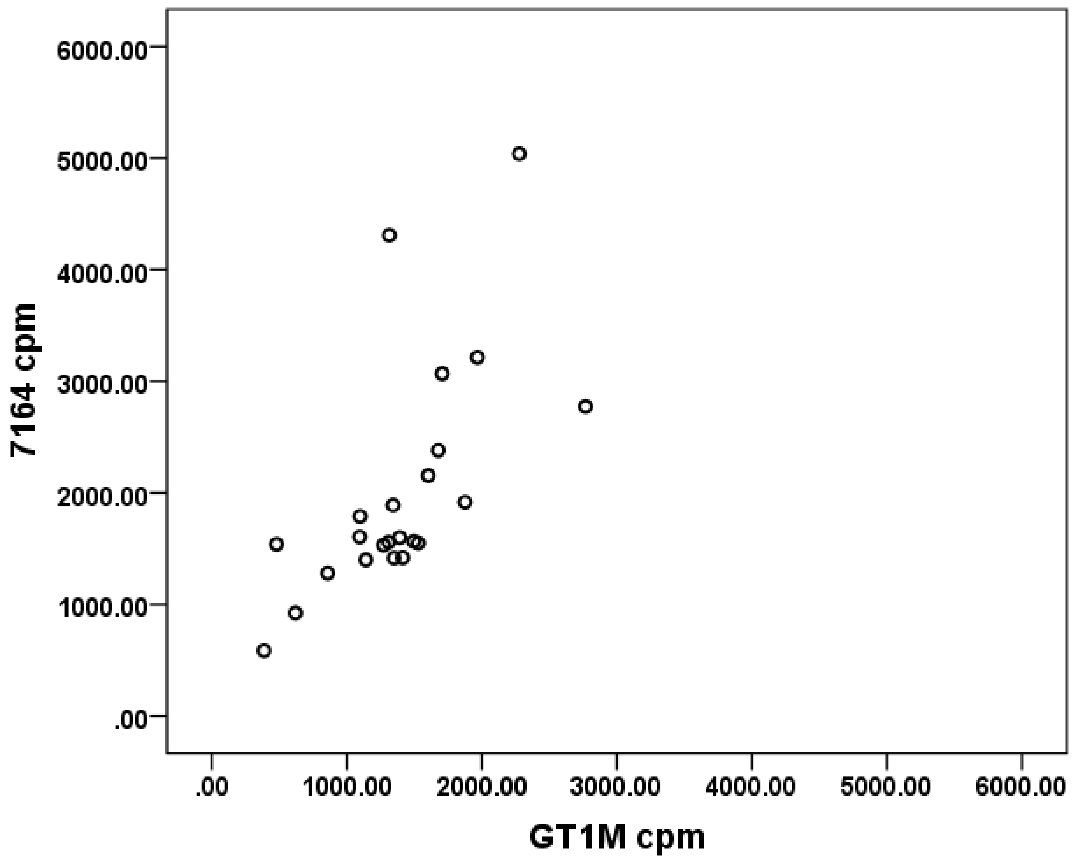
Appendix Table VII.i: Results of the normality tests Chapter 7.

Variable	***df***	**Shapiro-Wilks (n<50) test**	***p*-value**
7164 cpm	23	0.96	0.534
GT1M cpm	23	0.83	0.001
GT1M Sed^{va}	23	0.99	0.998
GT1M Sed^{pu}	23	0.96	0.56
GT1M Sed^{s}	23	0.99	0.999
7164 Sed^{va}	23	0.94	0.143
7164 Sed^{pu}	23	0.98	0.928
7164 Sed^{s}	23	0.98	0.248
7164 Sed^{ev}	23	0.83	0.001
GT1M LPA^{va}	23	0.94	0.143
GT1M LPA^{pu}	23	0.91	0.03*
GT1M LPA^{s}	23	0.95	0.29
7164 LPA^{va}	23	0.94	0.204
7164 LPA^{pu}	23	0.84	0.002*
7164 LPA^{s}	23	0.90	0.020*
GT1M $MVPA^{va}$	23	0.95	0.257
GT1M: $MVPA^{pu}$	23	0.89	0.014*
GT1M: $MVPA^{s}$	23	0.92	0.080
GT1M $MVPA^{pa}$	23	0.93	0.104
7164 $MVPA^{va}$	23	0.89	0.019*
7164 $MVPA^{pu}$	23	0.90	0.026*
7164 $MVPA^{s}$	23	0.88	0.009*
7164 $MVPA^{pa}$	23	0.97	0.600
GT1M TPA^{va}	23	0.96	0.404
GT1M TPA^{pu}	23	0.92	0.067
GT1M TPA^{s}	23	0.95	0.32
GT1M TPA^{ev}	23	0.90	0.024

7164TPA[va]	23	0.93	0.123
7164: TPA[pu]	23	0.95	0.24
7164: TPA[s]	23	0.94	0.22
7164: TPA[ev]	23	0.91	0.039
GT1M TPA[ev-corr]	23	0.90	0.024
GT1M: Sed[s-corr]	23	0.99	0.99
GT1M: Sed[pu-corr]	23	0.96	0.515
GT1M: LPA[s-corr]	23	0.95	0.292
GT1M: LPA[pu-corr]	23	0.97	0.583
GT1M: MVPA[s-corr]	23	0.92	0.080
GT1M: MVPA[pu-corr]	23	0.89	0.014
GT1M: TPA[s-corr]	23	0.95	0.320
GT1M: TPA[s-corr]	23	0.95	0.26
CARS Sed	23	0.99	0.99
CARS: LPA	23	0.92	0.077
CARS: MVPA	23	0.92	0.069
CARS: TPA	23	0.92	0.075

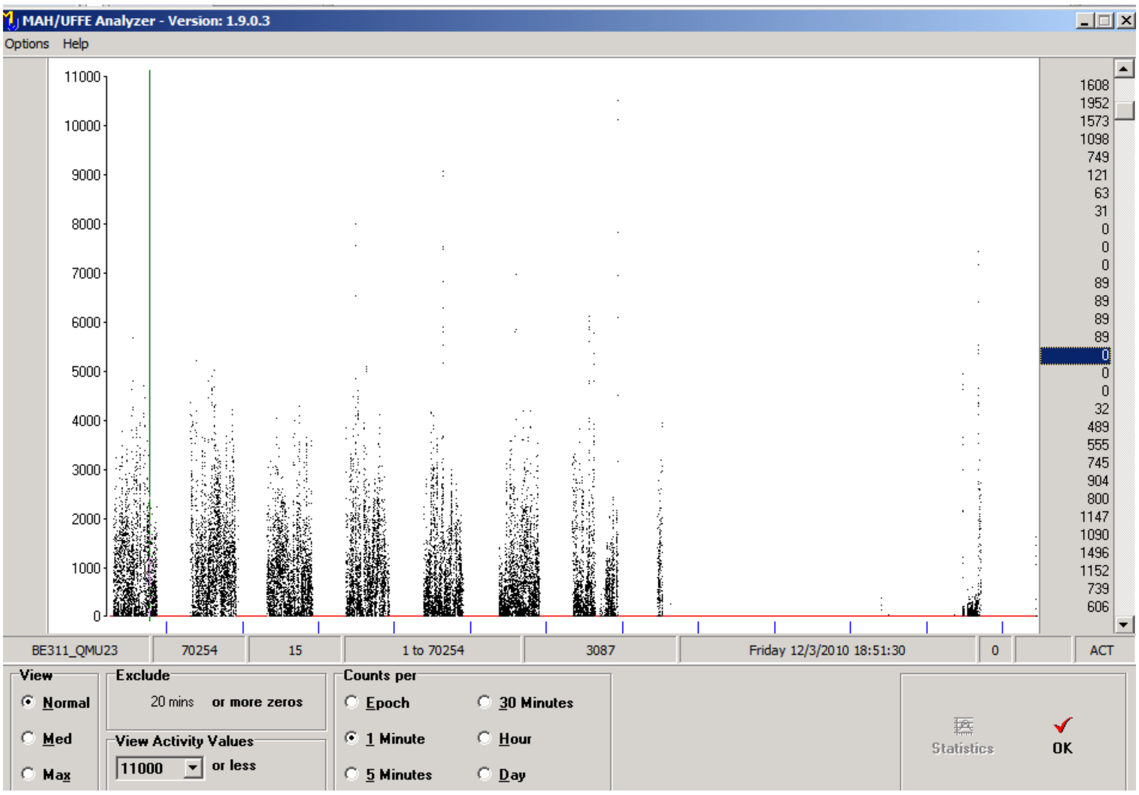
CARS: Children's Activity Rating Scale; [corr]*: GT1M data corrected GT1M/0.91; df: degrees of freedom; ev: Evenson et al.* (2008)*; LPA: light physical activity; MVPA: moderate-to-vigorous physical activity; pa: Pate et al.* (2006)*; pu: Puyau et al.* (2002)*; r: Reilly et al.* (2003)*; s: Sirard et al.* (2005)*; Sed: sedentary behaviour; TPA: total physical activity; va: Van Cauwenberghe et al.* (2011)

Appendix Figure VII.i: Scatter plot of the GT1M cpm and the 7164 cpm during free-living (n=23)

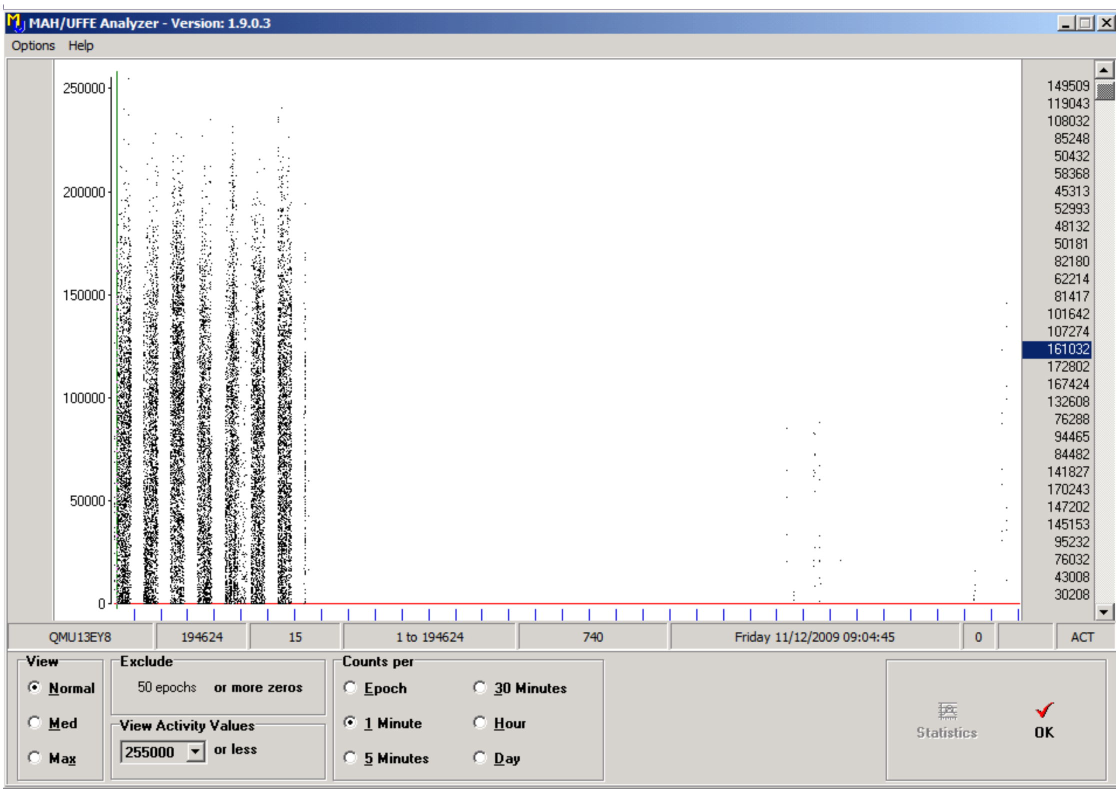


APPENDIX VIII: Output from MAHUffe programme and results of normality tests Chapter 8

Appendix Figure VIII.i: Output from MAHUffe programme with cpm < 20,000; left hand column presents the cpm.



Appendix Figure VIII.ii: Output from MAHUffe programme with cpm > 20,000; left hand column presents cpm. Suggesting errors with the activity the monitor.



Appendix Table VIII.i: Results of the normality tests Chapter 8.

Variable	**df**	**Kolmogorov-Smirnov (n>50) statistic**	***p*-value**
Registered wear time (weekend and weekdays)	112	0.096	0.013*
Registered time (weekdays)	112	0.115	0.001*
Registered time (weekend days)	103	0.059	0.200 [a]
Hours of wear time (weekend days)	201	0.094	0.001*
Hours of wear time (weekdays)	542	0.207	0.001*
cpm week days	103	0.104	0.008*
cpm weekend days	103	0.113	0.002*
cpm weekdays, males	60	0.150	0.002*
cpm weekend days, males	55	0.113	0.076
cpm weekdays, females	52	0.099	0.200[a]
cpm weekend days, females	48	0.174	0.001*
% time TPA weekdays 3 h	107	0.068	0.200 [a]
% time TPA weekdays 4 h	107	0.076	0.157
% time TPA weekdays 5 h	104	0.078	0.134
% time TPA weekdays 6 h	103	0.078	0.130
% time TPA weekdays 7 h	102	0.086	0.058
% time TPA weekdays 8 h	100	0.079	0.127
% time TPA weekdays 9 h	98	0.086	0.072
% time TPA weekdays 10 h	78	0.101	0.047
% time TPA weekend days 3 h	80	0.099	0.049
% time TPA weekend days 4 h	80	0.088	0.194
% time TPA weekend days 5 h	79	0.097	0.065
% time TPA weekend days 6 h	80	0.100	0.046
% time TPA weekend days 7 h	79	0.085	0.200[a]
% time TPA weekend days 8 h	77	0.079	0.200[a]
% time TPA weekend days 9 h	70	0.079	0.200[a]
% time TPA weekend days 10 h	58	0.091	0.200[a]

cpm: counts per minute; TPA: total physical activity

APPENDIX IX: Results of the normality tests Chapter 9

Appendix Table IX.i: Results of the normality tests Chapter 9.

Variable	***df***	**Kolmogorov-Smirnov (n>50) test**	***p*-value**
Sed^{ev}	104	0.099	0.013
Sed^{pu}	104	0.071	.200
Sed^{s}	104	0.061	.200
TPA^{ev}	104	0.039	.200
TPA^{pu}	104	0.063	.200
TPA^{s}	104	0.048	.200

df: degrees of freedom; ev: Evenson et al. (2008)*; pu: Puyau et al.* (2002)*; r: Reilly et al.* (2003)*; Sed: sedentary behaviour; s: Sirard et al.* (2005)*; TPA: total physical activity.*

APPENDIX X: Publications arising from this thesis